



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 142090**

**TO: Phuong Bui**  
**Location: REM/2A15/2C18**  
**Art Unit: 1638**  
**Thursday, January 13, 2005**

**Case Serial Number: 10/629953**

**From: Toby Port**  
**Location: Biotech-Chem Library**  
**REM-1A59**  
**Phone: 571-272-2523**

**toby.port@uspto.gov**

### **Search Notes**

Examiner Bui,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Toby Port

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 11, 2005, 13:59:30 ; Search time 3026 Seconds  
(without alignments)  
9876.765 Million cell updates/sec

Title: US-10-629-953-3  
Perfect score: 632  
Sequence: 1 ctctctcaagactccacaag.....taaaaaaaaaaaaaaaaaaaaaa 632

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	610.8	96.6	691	8 AK072440	AK072440 Oryza sat
2	349.2	55.3	403	8 OSU16255	U16255 Oryza sativ
3	324.4	51.3	145419	8 AP004876	AP004876 Oryza sat
4	182.2	28.8	289	6 AR246100	AR246100 Sequence
5	149.6	23.7	263	6 AR246253	AR246253 Sequence
6	145.6	23.0	286	6 AR231557	AR231557 Sequence
7	96	15.2	324	8 BT004745	BT004745 Arabidops
8	96	15.2	605	8 AF053302	AF053302 Arabidops
9	68.2	10.8	913	8 AB050390	AB050390 Brassica
10	65.4	10.3	22758	8 AP004508	AP004508 Lotus cor
11	64.4	10.2	923	8 AK104192	AK104192 Oryza sat
12	64.4	10.2	945	8 AK070545	AK070545 Oryza sat
13	64.4	10.2	1141	8 AK059936	AK059936 Oryza sat
14	62.4	9.9	82896	8 AL590346	AL590346 Arabidops
15	62.4	9.9	103534	8 ATT2K12	ATT2K12 Arabidops
16	61	9.7	498	8 BT005847	BT005847 Arabidops
17	61	9.7	730	8 AF053303	AF053303 Arabidops
18	61	9.7	757	8 AY085896	AY085896 Arabidops
19	55.4	8.8	7218	6 I66494	I66494 Sequence 14

20	55.2	8.7	2000	6	AX655393	AX655393 Sequence
21	54.4	8.6	313846	1	AE017234	AE017234 Mycobacte
22	53	8.4	1293	6	BD180561	BD180561 Highly th
23	52.4	8.3	333	8	BT004696	BT004696 Arabidops
24	52.4	8.3	538	8	AK117904	AK117904 Arabidops
25	52.4	8.3	1074	6	AR105098	AR105098 Sequence
26	52.4	8.3	1074	8	AF244702	AF244702 Zea mays
27	51.4	8.1	273285	1	AE017304	AE017304 Thermus t
28	51.2	8.1	70162	2	AC026568	AC026568 Homo sapi
29	51	8.1	4725	1	AF142435	AF142435 Frankia s
30	51	8.1	314100	1	SC0939106	SC0939106 Streptomy
31	50.2	7.9	110000	8	CR322131_34	Continuation (35 o
32	49.2	7.8	795	8	AY224434	AY224434 Oryza sat
33	49.2	7.8	795	8	AY224524	AY224524 Oryza sat
34	49.2	7.8	1152	8	AK101599	AK101599 Oryza sat
35	49.2	7.8	131049	8	AP005391	AP005391 Oryza sat
36	49.2	7.8	160089	8	AP004643	AP004643 Oryza sat
37	48.8	7.7	3443	8	AK120294	AK120294 Oryza sat
38	48.8	7.7	135216	8	AC069145	AC069145 Oryza sat
39	48.8	7.7	144724	8	AC078948	AC078948 Oryza sat
40	48.8	7.7	300029	8	AE017104	AE017104 Oryza sat
41	48.4	7.7	348257	1	BX640425	BX640425 Bordetell
42	48	7.6	100670	2	AP004682	AP004682 Oryza sat
43	48	7.6	148207	2	AP004697	AP004697 Oryza sat
44	48	7.6	170701	8	AP003510	AP003510 Oryza sat
45	48	7.6	301846	1	AE016913	AE016913 Chromobac

#### ALIGNMENTS

AK072440 691 bp mRNA linear PLN 24-JUL-2003  
Oryza sativa (japonica cultivar-group) cDNA clone:J023108D18, full  
insert sequence.

AK072440 GI:32982463

FLI CDNA; CAP trapper.

Oryza sativa (japonica cultivar-group)

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.

#### REFERENCE

1

AUTHORS

The Rice Full-length cDNA Consortium, National Institute of  
Agrobiological Sciences Rice Full-length cDNA Project Team,  
Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,  
Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Hotta,I.,  
Kojima,K., Namiki,T., Ohneda,E., Yabagi,W., Suzuki,K., Li,C.,  
Ohtsuki,K., Shishiki,T., Foundation of Advancement of International  
Science Genome Sequencing & Analysis Group: Ohtsuki,K., Murakami,K.,  
Tada,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,  
Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,  
Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Nikiura,J.,  
Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,  
Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN,  
Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,  
Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,  
Kagawa,R., Kondo,S., Konno,H., Miyazaki,A., Otsu,T., Ota,Y.,  
Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,  
Yoshino,M. and Hayashizaki,Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from  
japonica rice

Science 301 (5631), 376-379 (2003)

22752273

12869764

2 (bases 1 to 691)

REFERENCE

AUTHORS

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,  
Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,  
Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,  
Hori,F., Hotta,I., Iida,Y., Ikeda,R., Imamura,K.,  
Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,

Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otsu, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission  
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.  
URL: <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers  
1. .691  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J023108D18"

FEATURES  
Source  
1. .691  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J023108D18"

ORIGIN  
Query Match 96.8%; Score 610.8; DB 8; Length 691;  
Best Local Similarity 99.7%; Pred. No. 8e-103;  
Matches 612; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
1 CTCTCCAGACTCCACAGTCCAGAGGAGAGAGGTGTGAGAGAGATGTGGCGGAAGGGG 60  
42 CTCTCCAGACTCCACAGTCCAGAGGAGAGAGGTGTGAGAGAGATGTGGCGGAAGGGG 101  
61 AACAAAGCGGTTCCG 120  
102 AACAAAGCGGTTCCG 161  
121 GGCGCCCTCCAGAGCG 180

Db 162 GGGCCCTCCGAGAGCG 221  
QY 181 GTGGCGGTCCGAGACCTTGGAAACCGCAAGGTCTGTCGACATCCCGGAGTTCTACGGAAG 240  
Db 222 GTGGCGGTCCGAGACCTTGGAAACCGCAAGGTCTGTCGACATCCCGGAGTTCTACGGAAG 281  
QY 241 GACGGCAAGACCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300  
Db 282 GACGGCAAGACCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 341  
QY 301 CTGAGGGCAAAATATCAAGCTATAGATGAGGCGCATCAAGGAGAAATCGCTGATCGGAGCCC 360  
Db 342 CTGAGGGCAAAATATCAAGCTATAGATGAGGCGCATCAAGGAGAAATCGCTGATCGGAGCCC 401  
QY 361 ATTCTCTTGTGATGCAAGTAGACTAAGCCCTAGCTCTGTCTTTTATGACCTCCGAGGAATAT 420  
Db 402 ATTCTCTTGTGATGCAAGTAGACTAAGCCCTAGCTCTGTCTTTTATGACCTCCGAGGAATAT 461  
QY 421 TGACATTTTGGTATGATTAATCTTTATTAACCTGTCTAGTAATATCAAGTAATGTTTTAGT 480  
Db 462 TGACATTTTGGTATGATTAATCTTTATTAACCTGTCTAGTAATATCAAGTAATGTTTTAGT 521  
QY 481 TTGGTCATCGCGCGGAAATGTGCGCTTGTCTGTATTTGTCGATTTGTCGATTTGTCG 540  
Db 522 TTGGTCATCGCGCGGAAATGTGCGCTTGTCTGTATTTGTCGATTTGTCGATTTGTCG 581  
QY 541 CTAATGTGCTCTCCCTCATTTTGTAGTCGTTGAGGAGGAGTCCATCTGTTAAATGTTGGTT 600  
Db 582 CTAATGTGCTCTCCCTCATTTTGTAGTCGTTGAGGAGGAGTCCATCTGTTAAATGTTGGTT 641  
QY 601 GGTTCCTGCGGTGA 614  
Db 642 GGTTCCTGCGGTGA 655

RESULT 2  
OSU16255  
LOCUS  
DEFINITION  
Oryza sativa clone ST1 sucrose-regulated mRNA, 3'-end sequence.  
ACCESSION  
U16255  
VERSION  
U16255.1 GI:576770  
KEYWORDS  
Oryza sativa  
SOURCE  
Oryza sativa  
ORGANISM  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
REFERENCE  
1 (bases 1 to 403)  
Tseng, T.C., Tsai, T.H., Lue, M.Y. and Lee, H.T.  
Identification of sucrose-regulated genes in cultured rice cells  
using mRNA differential display  
Gene 161 (2), 179-182 (1995)  
JOURNAL  
MEDLINE  
95394352  
PUBMED  
7665075  
REFERENCE  
2 (bases 1 to 403)  
Lee, H.-t.  
Direct Submission  
Submitted (24-OCT-1994) Hung-tu Lee, National Tsing-Hua University,  
Life Science, No. 101, Sec. 2, Ken-Fu Rd., Hsinchu, Taiwan, R.O.C.,  
300  
COMMENT  
On Nov 28, 1994 this sequence version replaced gi:571441.  
FEATURES  
Location/Qualifiers  
1. .403  
/organism="Oryza sativa"  
/mol\_type="mRNA"  
/strains="Taiwan 5"  
/db\_xref="taxon:4530"  
/clone="S11"  
/cell\_type="suspension cultured cell"  
misc\_feature  
1. .403  
/note="induced by sucrose"  
/evidence="experimental"  
polyA\_site  
403



ORIGIN /note="11 A nucleotides"

Query Match 55.3%; Score 349.2; DB 8; Length 403;  
Best Local Similarity 99.28; Pred. No. 1.5e-54;  
Matches 351; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 261 CCGCAAGGTATACAGTCCCAATGGATCAGTGAAGATCTAGGAGCAATATCAAGC 320  
DB 2 CCGCAAGGTATACAGTCCCAATGGATCAGTGAAGATCTAGGAGCAATATCAAGC 61

QY 321 TATAGATGAGCCATCAAGAGATCGTGTATCGAGCCCATCTCTGTGATGCAAGTA 380  
DB 62 TATAGATGAGCCATCAAGAGATCGTGTATCGAGCCCATCTCTGTGATGCAAGTA 121

QY 381 GACTAAGCCTACGCTCTCTTTATGATCTCGAGGAATATTCGACTTTTGGTATGGTAAT 440  
DB 122 GACTAAGCCTACGCTCTCTTTATGATCTCGAGGAATATTCGACTTTTGGTATGGTAAT 181

QY 441 CTTTATTACCTGCTCTAGTAAATATCAAGTAATGTTTGGTGTATGCGGAAATG 500  
DB 182 CTTTATTACCTGCTCTAGTAAATATCAAGTAATGTTTGGTGTATGCGGAAATG 241

QY 501 TGGCTTGTCTGTTATGTCGCAAGAGCAGCTATGTGCTTAATGCTCTCTCATTT 560  
DB 242 TGGCTTGTCTGTTATGTCGCAAGAGCAGCTATGTGCTTAATGCTCTCTCATTT 301

QY 561 TGTACGTTGGTAAAGAGGTCATCTGTTAATGTTGGTGGTTTCGCGTGTA 614  
DB 302 TGTACGTTGGTAAAGAGGTCATCTGTTAATGTTGGTGGTTTCGCGTGTA 355

RESULT 3  
AP004876 LOCUS AP004876 145419 bp DNA linear PLN 28-JUL-2004  
DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, Chromosome 2,  
PAC clone:P0470G10.  
ACCESSION AP004876  
VERSION AP004876.3 GI:46390400  
KEYWORDS Oryza sativa (japonica cultivar-group)  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Erihartoideae; Oryzeae; Oryza.

REFERENCE  
1 Sasaki, T., Matsumoto, T. and Yamamoto, K.  
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, PAC  
clone:P0470G10  
Published Only in Database (2002)  
2 (bases 1 to 145419)  
Sasaki, T., Matsumoto, T. and Yamamoto, K.  
Direct Submission  
Submitted (20-MAR-2002) Takuji Sasaki, National Institute of  
Agrobiological Sciences, Rice Genome Research Program, Kannondai  
2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,  
Tel:81-298-38-7441, Fax:81-298-38-7468)  
On Apr 14, 2004 this sequence version replaced gi:34850224.  
Genes were predicted from the integrated results of the following:  
GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH  
(http://www.softberry.com/), GeneMark.hmm  
(http://opal.biology.gatech.edu/GeneMark/), GlimmerM  
(http://www.tigr.org/tdb/glimmer/glmr\_form.html), RiceHMM  
(http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor  
(http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4  
(http://globin.cse.psu.edu/html/doc/sim4.html), gap2  
(http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The  
genomic sequence was searched against NCBI NonRedundant Protein  
database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA  
sequence database at RGP or DDBJ. Protein homologies of the coding  
regions were searched against NCBI NonRedundant Protein database  
with BLASTP. ESTs represent the identified cDNA sequences using

BLASTN with the corresponding DDBJ accession no. and RGP clone ID.  
Full-length cDNAs represent the identified cDNA sequences using  
BLASTN with the corresponding DDBJ accession no.  
A gene with identity or significant homology to a protein is  
classified based on the protein name to indicate the homology level  
such as same name, 'putative-' and '-like protein'. A gene without  
significant homology to any protein but with full-length cDNA or  
EST homology (covering almost the entire length of partial  
sequence) is classified as an 'unknown' protein. A gene predicted  
by two or more gene prediction programs is classified as a  
'hypothetical' protein according to IRGSP standard. A gene  
predicted by a single gene prediction program is also classified as  
a probable 'hypothetical' protein and is included as a  
miscellaneous feature of the sequence.  
The orientation of the sequence is from SP6 to T7 of the PAC clone.  
This sequence of P0470G10 clone has an overlap with OJ1004.H01  
(DDBJ: AP004038) clone at 5' end and with OJ1008.F08 (DDBJ:  
AP004017) clone at 3' end. Detailed information on overlap and  
assembly quality together with annotation of this entry is  
available at <http://rgp.dna.affrc.go.jp/GenomeSeq.html>.

FEATURES  
Location/Qualifiers  
1..145419  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="2"  
/clone="P0470G10"  
complement(340..3372)  
/gene="P0470G10.1"  
complement(340..3372)  
/gene="P0470G10.1"  
/notes="probably inactive due to 5' exon missing in CDS  
pseudogene, receptor protein kinase"  
/pseudo  
complement(4490..5523)  
/gene="P0470G10.2"  
complement(4490..5523)  
/gene="P0470G10.2"  
/notes="probably inactive due to 3' exon missing in CDS  
pseudogene, Epstein-Barr virus EBNA-1-like protein"  
/pseudo  
complement(7419..8108)  
/gene="P0470G10.3"  
complement(7419..8108)  
/gene="P0470G10.3"  
/notes="probably inactive due to 3' exon missing in CDS  
pseudogene, Hcr2-0A"  
/pseudo  
complement(8546..10212)  
/gene="P0470G10.4"  
complement(join(<8546..8551,9081..9188,10147..>10212))  
/gene="P0470G10.4"  
/notes="start and end point are not identified"  
complement(join(8546..8551,9081..9188,10147..10212))  
/gene="P0470G10.4"  
/notes="predicted by GENSCAN etc."  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAD33295.1"  
/db\_xref="GI:50725764"  
/translation="MARQARAVVALVRIEHLNKGAWRRRLILAIQDDVELLLLS  
LLVLVGGTGGPHLQ"  
complement(11482..14658)  
/gene="P0470G10.5"  
complement(join(<11482..11634,11714..11941,12441..12545,  
12620..12697,12780..12854,13509..13595,13688..13755,  
13868..13955,14055..14155,14265..14385,14434..14479,  
14585..>14658))  
/gene="P0470G10.5"  
/notes="start and end point are not identified"  
complement(join(11482..11634,11714..11941,12441..12545,  
12620..12697,12780..12854,13509..13595,13688..13755,

```
13868..13955,14055..14155,14265..14385,14434..14479,
14585..14658))
/genes="P0470G10.5"
/note="contains full-length cDNA(s): AK100016,AK058244"
/codon_start=1
/product="putative thylakoid-bound ascorbate peroxidase"
/protein_id="BAD33296.1"
/db_xref="GI:50725765"
/translations="MRGTGRRRRRGGAGAPARGRWGSGVTPAPSQLSHPYBOVRLG
WDSQYDNIEPWRGAGDGLSRFDALSHGANGLNALKLIQPIKDKYPGITVA
DLFOLASATAIEAGGPKPKPMKYGRVDVTAACQPEGRLPDAGRPVPADHLREVFYR
MGJDDIEVALSGAHTLGRSPDRSGWGPETKYTKDGGEGPGQSWTVEWLKFDNSY
FKDIKEQRDLVLTDAALFDPSPFKVYAEKIDQAFKFDYAEAHAKLSDLGAK
FDPPEGFLDDEPAVEKDEPEAPAPAAAPPPEPVEEKEAEPFPVPTVGAAVASSP
ADNNGGAAPQEPBPFAAKYSYGGKELSDSMKQIRAEYEGFSGSPDKPLQSNYFLNM
LIGGLAFTSLGS"
16256..16777
/genes="P0470G10.6"
<16256..>16777
/genes="P0470G10.6"
/note="start and end point are not identified"
16256..16777
/genes="P0470G10.6"
/note="predicted by GeneMark.hmm etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD33297.1"
/db_xref="GI:50725766"
/translations="MKTPSPREATEAPPPTSVKVSKRTETRLSPGRAARGVETAQ
QRPEGDARNANPTKAGLGFHTLSVACISTADTPLLHHLNLRWRVGTTPAPPAPAS
LNPFPFAPVHRVALCTSPASTSASLASCAGHSMRHAGLRHWPRLSPAPPSVAATS
QTRLASLHTGCAS"
complement(18792..19088)
/genes="P0470G10.7"
complement(18792..19088)
/genes="P0470G10.7"
/note="hypothetical ORF
predicted by GlimmerM
this category is not included in IRGSP standard"
complement(20341..21044)
/genes="P0470G10.8"
complement(<20341..>21044)
/genes="P0470G10.8"
/note="supported by full-length cDNA(s): AK110958"
complement(20341..21044)
/genes="P0470G10.8"
/note="contains full-length cDNA(s): AK110958
non-coding transcript
probably inactive due to including stop codon(s) in CDS"
22104..25416
/genes="P0470G10.9"
join(<22104..22246,23442..23556,24206..24344,24782..24813,
24943..25029,25192..>25416)
/genes="P0470G10.9"
/note="start and end point are not identified"
join(22104..22246,23442..23556,24206..24344,24782..24813,
24943..25029,25192..>25416)
/genes="P0470G10.9"
/note="contains EST(s): C28326(C60709)"
/codon_start=1
/product="putative tetratricopeptide repeat
(TPR)-containing protein"
/protein_id="BAD33298.1"
/db_xref="GI:50725767"
/translations="MRGFLMHQLLSDSSDDDELILAAALIAHQHYDIDNAPRRRG
SVASNDTEKEBTRVTTDCPTKPSDNABSLKHQNLRAEKGKHEALGRWEAALT
LMPNNAITLHQKQAILLEKGDRAWALTAATRA TELDLPWPGVITLGRALQNFGEPS
AIIISFDKALAIEPDNEAKSDRTAARLVKKRQLHSSGLSANKRRFTVGEDSEKED
DSOMEDVEKCKEDEERDSDLENVEK"
27139..29341
/genes="P0470G10.10"
join(27139..27341,27451..27556,28514..28536,28983..29341)
/genes="P0470G10.10"
```

```
CDS
/note="supported by full-length cDNA(s): AK072440"
join(27225..27341,27451..27556,28514..28536,28983..29042)
/genes="P0470G10.10"
/note="contains EST(s): AU082888(C53889),AU068340(C20066)
contains full-length cDNA(s): AK072440"
/codon_start=1
/product="transcriptional coactivator p15 (PC4) family
protein-like"
/protein_id="BAD33299.1"
/db_xref="GI:50725768"
/translations="WRRGNKFFGGGEPAAKRAAGDDPSADDDDIIVAOISKNR
RVAVRWNGKVVDIREFYEKQKYLPRKGIQLPMDQWKLRLDNLIKADIEAIKENA"
complement(30379..31415)
/genes="P0470G10.11"
complement(join(<30379..30401,31091..>31415))

Query Match 51.3%; Score 324.4; DB 8; Length 145419;
Best Local Similarity 99.7%; Pred. No. 5.4e-50;
Matches 325; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 289 CAGTGAAGATCTAGGAGCAATATCAAAGCTATAGATGAGGCCATCAAGGAGATGCG 348
DB 28980 CAGTGAAGATCTAGGAGCAATATCAAAGCTATAGATGAGGCCATCAAGGAGATGCG 29039
QY 349 TGATCGAGCCCATCTCTTGTGATGCAAGTAGACTAAGCCTACGCTCTCTTTTATGAC 408
DB 29040 TGATCGAGCCCATCTCTTGTGATGCAAGTAGACTAAGCCTACGCTCTCTTTTATGAC 29099
QY 409 TCCGAGGATATGCACTTTGGTATGTAATCTTTATACCTGCTAGTATATCAAG 468
DB 29100 TCCGAGGATATGCACTTTGGTATGTAATCTTTATACCTGCTAGTATATCAAG 29159
QY 469 TAATGTTTGTGTCATGCGCGAAATGTGGCTTGTCTGGTATTGTGCAAGAA 528
DB 29160 TAATGTTTGTGTCATGCGCGAAATGTGGCTTGTCTGGTATTGTGCAAGAA 29219
QY 529 GCAGCTATGTGCTTAATGTCTCTCTCAATTTGTAGCTTGGGTAGGAGGTCCATCTGT 588
DB 29220 GCAGCTATGTGCTTAATGTCTCTCTCAATTTGTAGCTTGGGTAGGAGGTCCATCTGT 29279
QY 589 TAATGTTTGTGTTTCTGCGGTGA 614
DB 29280 TAATGTTTGTGTTTCTGCGGTGA 29305

RESULT 4.
AR246100 AR246100 289 bp DNA linear PAT 20-DEC-2002
LOCUS Sequence 1459 from patent US 6476212.
DEFINITION AR246100
ACCESSION AR246100
VERSION AR246100.1 GI:27293974
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 289)
AUTHORS Laluguri,R.V., Ito,L.Y. and Sherman,B.K.
TITLE Polynucleotides and polypeptides derived from corn ear
JOURNAL Patent: US 6476212-A 1459 05-NOV-2002;
FEATURES Location/Qualifiers
source 1..289
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 28.8%; Score 182.2; DB 6; Length 289;
Best Local Similarity 79.6%; Pred. No. 9.4e-24;
Matches 230; Conservative 0; Mismatches 53; Indels 6; Gaps 1;

QY 41 GAAGAATGTGCGGAAAGGGAACAAGCGTTCGCGCGCGCGCGC-----GAGCGCGCG 94
DB 1 GGAAGATGTGCGGGAAGGAAGAAAGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCG 60
```

QY 95 CCAAGCGCGTGCCTCGCGGGGACGACGGGCGCTCCGAGAGCGCGGACGATATCGTCG 154  
Db 61 CCAAGCGCGCGCGCGGAGGAGATGCGCCCTCCGAACTCCGTAAGACGGTACCGTCG 120  
QY 155 TCGCCAGATATCGAAGAACAGAGGAGGTGCGGTGCGGACCTCGAAGCGCAAGGTCTCG 214  
Db 121 TAGCCGAGATATCGAAGAACAGAGAGGTGTCGTTAGGAGCTGGAAGGCGAGGTCTTCG 180  
QY 215 TCACATCCCGAGTTCTACGAGAACGCGCAAGACCTCCCGCGCGCAAGGTATAC 274  
Db 181 TCAGCTTACCGAGTTCTACTTCAAGGACGCGAAGACTCTCCACCACCGCAAGGTATAT 240  
QY 275 AGCTCCCAATGGATCACTAGTGAAGTACTGAGGACATATCAAGCTAT 323  
Db 241 CACTCCAAATAGATCAGTGAAGATATTGAAGGACATCAAGGCCAT 289

RESULT 5  
AR246253  
LOCUS AR246253 263 bp DNA linear PAT 20-DEC-2002  
DEFINITION Sequence 1612 from patent US 6476212.  
ACCESSION AR246253  
VERSION AR246253.1 GI:27294127  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 263)  
AUTHORS Lalgudi,R.V., Ito,L.Y. and Sherman,B.K.  
TITLE Polynucleotides and polypeptides derived from corn ear  
JOURNAL Patent: US 6476212-A 1612 05-NOV-2002;  
FEATURES  
source Location/Qualifiers  
1..263  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 23.7%; Score 149.6; DB 6; Length 263;  
Best Local Similarity 78.0%; Pred. No. 9.7e-18;  
Matches 206; Conservative 0; Mismatches 51; Indels 7; Gaps 2;  
QY 41 GAAGATGTGCGGAGGGAACAGCGGTTCGCGCGCGCGC-----GAGCGCGCG 94  
Db 1 GGAAGATGTGCGGAGGGAAGAGCGGTTCGCGCGCGCGCGCGCGCGCGCG 60  
QY 95 CCAAGCGCGTGCCTCGCGGGACGACGGGCGCTCCGAGAGCGCGGACGATATCGTCG 154  
Db 61 CCAAGCGCGCGCGCGGAGGAGCGATGCGCCCTCCGATCCGCTGAGACGGTACCGTCG 120  
QY 155 TCGCCAGATATCGAAGAACAGAGGAGGTGCGGTGCGGACCTCGAAGCGCAAGGTCTCG 214  
Db 121 TAGCCGAGATATCGAAGAACAGAGAGGTGTCGTTAGGAGCTGGAAGGCGAGGTCTTCG 180  
QY 215 TCACATCCCGAGTTCTACGAGAACGCGCAAGACCTCCCGCGCGCAAGGTATAC 274  
Db 181 NCAGCTTACCGAGTTCTACTTCAAGGACGCGAAGACTCTCCACCACCGCAAGGTATA- 239  
QY 275 AGCTCCCAATGGATCAGTGAAGA 298  
Db 240 TCATCCAAATAGATCAGTGAAGA 263

RESULT 6  
AR251557  
LOCUS AR251557 286 bp DNA linear PAT 20-DEC-2002  
DEFINITION Sequence 6916 from patent US 6476212.  
ACCESSION AR251557  
VERSION AR251557.1 GI:27299431  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 286)

AUTHORS Lalgudi,R.V., Ito,L.Y. and Sherman,B.K.  
TITLE Polynucleotides and polypeptides derived from corn ear  
JOURNAL Patent: US 6476212-A 6916 05-NOV-2002;  
FEATURES  
source Location/Qualifiers  
1..286  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 23.0%; Score 145.6; DB 6; Length 286;  
Best Local Similarity 78.4%; Pred. No. 5.3e-17;  
Matches 189; Conservative 0; Mismatches 45; Indels 7; Gaps 1;  
QY 40 AGAAGAATGTGCGGAGGGAACAAGCGGTTC-----GGCGGCGCGCGGAGCGCGC 92  
Db 44 AGCAAGATGTGCGGAGGGAAGAAAGCGTTTCGCGCGGTGGCGGCGAGCGGTGAGCCGC 103  
QY 93 GGCNAAGCGCGTGCCTCGCGGGGACGACGGGCGCTCCGAGAGCGCGGACGATATCGT 152  
Db 104 GGCNAAGCGCGAGCGCGGAGGAGCGATGCGCCCTCCGAATCCGTAAGACGGTACCGT 163  
QY 153 CGTCGCGCCAGATATCGAAGAACAGGAGGCGTGGCGGTGCGGACCTGGAAACGGCAAGGTCT 212  
Db 164 CGTAGCGGAGATATCGAAGAACAGAGGTGTCGTTAGGAGCTGGAAGCGCGGTCTT 223  
QY 213 CGTCGACATCCCGCGAGTTCTACGAGAACGAGCGCAAGACCTCCCGCGCGCGCAAGGTAT 272  
Db 224 CGTCGACTTACGCGAGTTCTACTTCAAGGAGCGCAAGACTCTCCCGCGCGCAAGGTAT 283  
QY 273 A 273  
Db 284 A 284

RESULT 7  
BT004745  
LOCUS BT004745 324 bp mRNA linear PLN 22-FEB-2003  
DEFINITION Arabidopsis thaliana At5g09250 gene, complete cds.  
ACCESSION BT004745  
VERSION BT004745.1 GI:28466804  
KEYWORDS FLI CDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
REFERENCE 1 (bases 1 to 324)  
AUTHORS Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Bowser,L., Carninci,P., Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Heuan,V.W., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.  
TITLE Arabidopsis ORF clones  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 324)  
AUTHORS Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Bowser,L., Carninci,P., Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Heuan,V.W., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.  
TITLE Direct Submission  
JOURNAL Submitted (22-FEB-2003) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN

Arabidopsis Full-length cDNA': Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PESC (SSP) Consortium members constructed and sequenced the PUN1 (ORF) clones using the RAP1 cDNAs: Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Chan, M.M., Chang, C.M., Dale, J.M., Huan, V.W., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

#### FEATURES

source

Location/Qualifiers

1..324

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="5"

/clone="U23341"

/scotopes="Columbia"

/notes="This clone is in pUNI 51"

1..324

/note="putative transcriptional co-activator (KIWI)"

/codon\_start=1

/product="At5g09250"

/protein\_id="AAO44011.1"

/db\_xref="GI:28466805"

/translation="MSSRGKRDDEVRASDSETHAPAKVAKPADDSDQSDDIIVVC

NISKNRVSVRNWNGKIWDIREFYVDGKTLPGKKGISLSDVQWTLNRHAEDIEKA

LSDSL"

#### ORIGIN

Query Match 15.2%; Score 96; DB 8; Length 324;  
Best Local Similarity 66.3%; Pred. No. 7.5e-08;  
Matches 138; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 142 GACGATATCGTCTGCCCGCAGATATCGAAGAACAGAGGGTGGCGGTGGGACCTGGAAC 201  
|||||  
DB 115 GACGATATCGTCTGCGCAACATATCTAAGATAGGAGAGTCTCTGTAAGGAATTGGAAC 174  
|||||

QY 202 GCGAAGTCTGCTCGACATCCCGAGTCTACGAGAGCGGACGACCTCCCGGC 261  
|||||

DB 175 GCGAAGATTGGATTGACATTCGTGAGTCTATGTCAAGACGGAAGACTTTGCTGGC 234  
|||||

QY 262 CGCAAGGTATACAGTCTCCCAATGGATCAGTGAAGATCTGAGGACCAATATCAAGCT 321  
|||||

DB 235 AAGAAAGGTATCTCTTAAGCGTGGATCAGTGAACACTTTTCGGAAACCGCAGAGGAT 294  
|||||

QY 322 ATAGATGAGGCCATCAAGGAGATGCGT 349  
|||||

DB 295 ATCGAGAAGGCCCTCTCTGACCTTTCTT 322  
|||||

#### RESULT 8

AF053302

LOCUS

DEFINITION

Arabidopsis thaliana putative transcriptional co-activator (KIWI)

mRNA, complete cds.

AF053302

VERSION

AF053302.1 GI:2997683

KEYWORDS

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 605)

Cormack, R.S., Hahlbrock, K. and Somssich, I.E.

Isolation of putative plant transcriptional coactivators using a

#### JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (11-MAR-1998) Biochemistry, Max-Planck-Institut fuer

Zuechtungsforchung, Carl-von-Linne-Weg 10, Cologne 50829, Germany

Location/Qualifiers

source

1..605

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/cultivar="Columbia"

/db\_xref="taxon:3702"

/chromosome="V"

/map="between KG31 and nga249"

1..605

/gene="KIWI"

97..420

/gene="KIWI"

/codon\_start=1

/product="putative transcriptional co-activator"

/protein\_id="AAC08574.1"

/db\_xref="GI:2997684"

/translation="MSSRGKRDDEVRASDSETHAPAKVAKPADDSDQSDDIIVVC

NISKNRVSVRNWNGKIWDIREFYVDGKTLPGKKGISLSDVQWTLNRHAEDIEKA

LSDSL"

#### ORIGIN

Query Match 15.2%; Score 96; DB 8; Length 605;  
Best Local Similarity 66.3%; Pred. No. 7.5e-08;  
Matches 138; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 142 GACGATATCGTCTGCCCGCAGATATCGAAGAACAGAGGGTGGCGGTGGGACCTGGAAC 201  
|||||

DB 211 GACGATATCGTCTGCGCAACATATCTAAGATAGGAGAGTCTCTGTAAGGAATTGGAAC 270  
|||||

QY 202 GCGAAGTCTGCTCGACATCCCGAGTCTACGAGAAGCGGACGACCTCCCGGC 261  
|||||

DB 271 GCGAAGATTGGATTGACATTCGTGAGTCTATGTCAAGACGGAAGACTTTGCTGGC 330  
|||||

QY 262 CGCAAGGTATACAGTCTCCCAATGGATCAGTGAAGATATCTGAGGACCAATATCAAGCT 321  
|||||

DB 331 AAGAAAGGTATCTCTTAAGCGTGGATCAGTGAACACTTTTCGGAAACCGCAGAGGAT 390  
|||||

QY 322 ATAGATGAGGCCATCAAGGAGATGCGT 349  
|||||

DB 391 ATCGAGAAGGCCCTCTCTGACCTTTCTT 418  
|||||

#### RESULT 9

AB050390

LOCUS

DEFINITION

Brassica rapa BcKELP mRNA for putative transcriptional coactivator,

complete cds.

AB050390

VERSION

AB050390.1 GI:13699909

KEYWORDS

Brassica rapa

ORGANISM

Brassica rapa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1

Matsushita, Y., Deguchi, M., Yoda, M., Nishiguchi, M. and Nynuya, H.

The tomato mosaic tobamovirus movement protein interacts with a

putative transcriptional coactivator KELP

Mol. Cells 12 (1), 57-66 (2001)

JOURNAL

MEDLINE

PUBMED

2 (bases 1 to 913)

REFERENCE

AUTHORS

Matsushita, Y. and Nynuya, H.

modified two-hybrid system incorporating a GFP reporter gene  
Plant J. 14 (6), 685-692 (1998)

98346011

9681033

2 (bases 1 to 605)

Cormack, R.S., Hahlbrock, K. and Somssich, I.E.

Direct Submission

Submitted (11-MAR-1998) Biochemistry, Max-Planck-Institut fuer

Zuechtungsforchung, Carl-von-Linne-Weg 10, Cologne 50829, Germany

Location/Qualifiers

source

1..605

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/cultivar="Columbia"

/db\_xref="taxon:3702"

/chromosome="V"

/map="between KG31 and nga249"

1..605

/gene="KIWI"

97..420

/gene="KIWI"

/codon\_start=1

/product="putative transcriptional co-activator"

/protein\_id="AAC08574.1"

/db\_xref="GI:2997684"

/translation="MSSRGKRDDEVRASDSETHAPAKVAKPADDSDQSDDIIVVC

NISKNRVSVRNWNGKIWDIREFYVDGKTLPGKKGISLSDVQWTLNRHAEDIEKA

LSDSL"

**TITLE** Direct Submission  
**JOURNAL** Submitted (20-OCT-2000) Yasuhiko Matsushita, Tokyo University of Agriculture and Technology, Gene Research Center; 3-5-8 Saiwai-cho, Fuchu, Tokyo 183-8509, Japan (E-mail: ymatasu@cc.tuat.ac.jp, Tel: 81-42-367-5855, Fax: 81-42-367-5855)

**FEATURES** Location/Qualifiers  
 source  
 1. .913  
 /organism="Brassica rapa"  
 /mol\_type="mRNA"  
 /strain="S9 homozygote"  
 /db\_xref="taxon:3711"  
 /clone="MIP102 (=BCKELP)"  
 /clone\_lib="lambda GEX5 Brassica rapa flower bud cDNA"  
 /notes="vector: pGEX-Bc2"  
 1. .913  
 /gene="BCKELP"  
 14. .511  
 /gene="BCKELP"  
 /codon\_start=1  
 /product="putative transcriptional coactivator"  
 /protein\_id="BAB41214.1"  
 /db\_xref="GI:13699910"  
 /translation="MEESKAKIEETVREILKESDMTEFKNVNLASERLIGDILSD KSHKAPVGVKSLFEEVSKQOQDKEEERERAKENKGFDDGDLICRLS DRRVTIQFRGKSLVSIREYKKGKELPSSKGLSLTDQWSTFKXNIPATEAAVKK MESRV"  
 869. .874  
 /gene="BCKELP"  
 /evidence="not\_experimental"

**polya\_signal** 869. .874

**ORIGIN**

Query Match 10.8%; Score 68.2; DB 8; Length 913;  
 Best Local Similarity 57.9%; Pred. No. 0.01;  
 Matches 121; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

Qy 135 CGCCGACGACGATATCGTCTGCCAGATATCGAAGACAGAGGCGGTGGCGGAC 194  
 |||||  
 Db 286 CGATGACGGGATCTCATCTTTGCGAGGTGTCGGATAAGAGGAGTGACGATTCAGGA 345

Qy 195 CTGGAACGGCAAGTCTGCTGCACATCCGCGAGTTCTACGAGAAGCGCGCAAGCCCT 254  
 |||||  
 Db 346 GTTTAGAGGAGGAGTTGTTGTTCCATCAGAGATATTACAGAGAGCGCAAGAGCT 405

Qy 255 CCCC GCCC CCAAGGATATACGCTCCCAATGGATCAGTCAAGATACGAGGACAAATAT 314  
 |||||  
 Db 406 TCCTTCTTAAAGGAATAAGCTTAACAGACGACAACTGTCACCTTCAAGAAAAATAT 465

Qy 315 CAAGCTATAGTACGCCATCAGAGA 343  
 |||||  
 Db 466 TCAGCTATCGAAGTCTGTCAGAAAA 494

**RESULT 10**  
**AP004508/c**  
**LOCUS** 22758 bp DNA linear PLN 22-JUL-2003  
**DEFINITION** Lotus corniculatus var. japonicus genomic DNA, chromosome 1, clone: LjT03K03, TM0036c, complete sequence.

**ACCESSION** AP004508  
**VERSION** AP004508.1 GI:17736875  
**KEYWORDS** HTG.  
**SOURCE** Lotus corniculatus var. japonicus (Lotus japonicus)  
**ORGANISM** Lotus corniculatus var. japonicus  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.

**REFERENCE** 1  
**AUTHORS** Sato, S., Kaneko, T., Nakamura, Y., Asamizu, E., Kato, T. and Tabata, S.  
**TITLE** Structures and Mapping of Fifty-six TAC clones which cover the 5.4 Mb Regions of the Genome  
**Unpublished**  
 2 (bases 1 to 22758)

**AUTHORS** Nakamura, Y.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (13-DEC-2001) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: ynakam@kazusa.or.jp, URL: http://www.kazusa.or.jp, Tel: 81-438-52-3935, Fax: 81-438-52-3934)

**FEATURES** Location/Qualifiers  
 source  
 1. .22758  
 /organism="Lotus corniculatus var. japonicus"  
 /mol\_type="genomic DNA"  
 /variety="japonicus"  
 /db\_xref="taxon:34305"  
 /chromosome="1"  
 /clone="LjT03K03"  
 /clone\_lib="LjT library"  
 /note="TM0036c, a part of TAC clone: TM0036-synonym: Lotus japonicus"

**ORIGIN**

Query Match 10.3%; Score 65.4; DB 8; Length 22758;  
 Best Local Similarity 73.0%; Pred. No. 0.033;  
 Matches 84; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 158 CCAGATATCGAAGAACAGAGGCGGTGGCGGACCTTGGAACGCAAGGTCTGTCTG 217  
 |||||  
 Db 22653 CGCAGATTTTCAAGAACAGAGGCGTTTCTGTGAGGAACCTGGCAGGCGAGGATCGTGTCTG 272

Qy 218 ACATCCGCGAGTTCTACGAGAGGCGCAAGCCCTCCCGCCGCAAGGTAT 272  
 |||||  
 Db 22593 ACATTCGCGAATTTTATGTCAAAGCGCAAGAAATGCTTGGGAAGAAGGTTT 22539

**RESULT 11**  
**AK104192**  
**LOCUS** 923 bp mRNA linear PLN 24-JUL-2003  
**DEFINITION** Oryza sativa (japonica cultivar-group) cDNA clone: 006-303-D08, full insert sequence.

**ACCESSION** AK104192  
**VERSION** AK104192.1 GI:32989401  
**KEYWORDS** FLI CDNA; oligo capping.  
**SOURCE** Oryza sativa (japonica cultivar-group)  
**ORGANISM** Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

**REFERENCE** 1  
**AUTHORS** The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Otonari, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurotsaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.  
 Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice  
 Science 301 (5631), 376-379 (2003)

**JOURNAL** MEDLINE 22752273  
**REFERENCE** 12869764  
 2 (bases 1 to 923)  
**AUTHORS** Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,





COMMENT	<p>This clone is one of the 28K full-length cDNA clones from japonica rice.</p> <p>URL : <a href="http://cdna01.dna.affrc.go.jp/cDNA/">http://cdna01.dna.affrc.go.jp/cDNA/</a></p> <p>NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yasaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing &amp; Analysis Group: Otomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurotsaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nariakawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN: Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, T., Kondo, S., Konno, K., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.</p> <p>Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice</p> <p>Science 301 (5631), 376-379 (2003)</p> <p>2 (bases 1 to 1141)</p>	
TITLE	japonica rice	
JOURNAL	22752273	
MEDLINE	12869764	
PUBMED	2 (bases 1 to 1141)	
REFERENCE	<p>Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, T., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurotsaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Negata, T., Nakamura, M., Namiki, T., Nariakawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yasaki, J., Yokomizo, S. and Yoshimura, A.</p>	
TITLE	Direct Submission	
JOURNAL	Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression: 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan [E-mail: <a href="mailto:skikuchi@nias.affrc.go.jp">skikuchi@nias.affrc.go.jp</a> , Tel: 81-29-838-7007, Fax: 81-29-838-7007]	
COMMENT	<p>This clone is one of the 28K full-length cDNA clones from japonica rice.</p> <p>URL : <a href="http://cdna01.dna.affrc.go.jp/cDNA/">http://cdna01.dna.affrc.go.jp/cDNA/</a></p> <p>NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yasaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.</p> <p>FAIS Genome Sequencing &amp; Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurotsaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariakawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.</p> <p>Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, T., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohta, Y., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.</p>	
FEATURES	<p>Location/Qualifiers</p> <p>1..945</p> <p>/organism="Oryza sativa (japonica cultivar-group)"</p> <p>/mol_type="mRNA"</p> <p>/cultivar="Nipponbare"</p> <p>/db_xref="taxon:39947"</p> <p>/clone="J023055M24"</p>	
source	10.2%; Score 64.4; DB 8; Length 945;	
ORIGIN	<p>Best Local Similarity 59.2%; Pred. No. 0.05;</p> <p>Matches 129; Conservative 0; Mismatches 86; Indels 3; Gaps 1;</p>	
QY	136	GCCGACGACGATATCTGTCGCCCGCCAGATATCGAAGACGAGGCGTGGCGGACC 195
Db	350	GACGAGGAGACCTCATCTCTCCCGCTTTCGCCGAGGAGGAGTCTTTCAGGAG 409
QY	196	TGGAACGCGCAAGTCTGTCGCACATCCGCGAGTTCTACGAGAAGCAGCGCAAGCCCTC 255
Db	410	TTCAAGGCGCAGAGCTTCTGTCATCCCGAGAGTACTTTCAGGACGCGCAGGAGCTG 469
QY	256	CCCGCGCGCAAGGATACAGTCCCAATGGATCATGTGAAGATACCTGAGGGCAATATC 315
Db	470	CCCG---CCAAAGGAGATAAGTTTGACAGTTGACAGTGGGAGCAATTCGTGATTCTGTG 526
QY	316	AAAGCTATAGTACGCCATCAGGAGGATGCGGTGATC 353
Db	527	CCAGCAATAGAGGATGCCATTAAAGAGCTTGGAGNATC 564
RESULT 13		
AK059936		
LOCUS	AK059936 1141 bp mRNA linear PLN 24-JUL-2003	
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:006-210-D10, full insert sequence.	
ACCESSION	AK059936	
VERSION	AK059936.1 GI:32969954	
KEYWORDS	FLI_CDN; oligo-capping.	
SOURCE	Oryza sativa (japonica cultivar-group)	
ORGANISM	Oryza sativa (japonica cultivar-group)	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; Poales; Poaceae; Oryzae; Ehrhartoideae; Oryzae; Oryza.	
REFERENCE	1	
AUTHORS	The Rice Full-length cDNA Consortium, National Institute of	

Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,  
Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,  
Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K.,  
Yasunishi,A. and Hayashizaki,Y.  
Location/Qualifiers  
1. .1141  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="006-210-D10"

ORIGIN  
Query Match 10.2%; Score 64.4; DB 8; Length 1141;  
Best Local Similarity 59.2%; Pred. No. 0.05;  
Matches 129; Conservative 0; Mismatches 86; Indels 3; Gaps 1;  
Qy 136 GCGCAGCAGCATATCTCTCGCCAGATATCGAAGAACAGGAGGTGCGGTGGACG 195  
Db 428 GACGAGGAGACCTCATCTCTCGCCCTTCGCCAGGAGGAGGTCTCTCAGGAG 487  
Qy 196 TGGACGGCAAGTCTGTCGACATCCGCGAGTTCTACGAGAGACGCGACCCCTC 255  
Db 488 TTCAAGGGCAAGCGCTTCTGTCATCCGAGAGTACTACTTCAAGGACGCGAAGGAGCTG 547  
Qy 256 CCGGCGCGCAAGGTATACAGCTCCCAATGGATCAGTGAAGATACCTGAGGACAAATATC 315  
Db 548 CCGG----CAAGGGATAGTTTGACAGTTGACAGTGGAGCATTTCTGTGTTCTGTG 604  
Qy 316 AAAGCTATAGTAGAGCCATCAAGGAGAGATCGGTGATC 353  
Db 605 CCAGCAATAGAGGATGCCATTAAGAGCTTGGAGATC 642

RESULT 14  
ATT2K12/c  
LOCUS  
DEFINITION Arabidopsis thaliana DNA chromosome 5, BAC clone T2K12 (ESSA project).  
ACCESSION AL590346  
VERSION 1  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
REFERENCE 1 (bases 1 to 82896)  
AUTHORS Bevan,M., Murphy,G., Ridley,P., Hudson,S., Bancroft,I., Mewes,H.W.,  
Rudd,S., Lemcke,K. and Mayer,K.F.X.  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 82896)  
AUTHORS EU Arabidopsis sequencing, project.  
TITLE Direct Submission  
JOURNAL Submitted (02-APR-2001) MIPS, at GSF/IBI, Ingolstaedter Landstr. 1,  
85764 Neuherberg, FRG, E-mail: schoofegsf.de, kmayer@gsf.de  
Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge  
Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,  
E-mail: michael.bevan@brc.ac.uk  
Information on performance of analysis and a more detailed  
annotation of this entry and other sequences of chromosomes 3, 4  
and 5 can be viewed at: <http://mips.gsf.de/proj/thal/>.

FEATURES  
source Location/Qualifiers  
1. .82896  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/variety="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="5"  
misc\_feature 1. .14902  
notes="overlap to PAC MAH20, EMBL:AB008697; please refer to this clone for annotation"  
15981. .19300  
gene

CDS  
/gene="At5g08660"  
join(15981..16289,16433..16746,16833..16925,17020..17110,  
17305..17398,17478..17575,17767..17836,18028..18131,  
18231..18347,18458..18534,18636..18680,18763..19300)  
/note="similarity to predicted proteins, Arabidopsis thaliana"  
/codon\_start=1  
/product="putative protein"  
/protein\_id="CAC35871.1"  
/db\_xref="GI:13548324"  
/db\_xref="trEMBL:Q9CSB1"  
/translation="MGSFCSKSLGINFGEYSGSVADGDREDFGHSPNGQTSILV  
PGMRQMLVKDQEQNLKDVFSFRERAEEDNFYDGIPTVTWAPSKIRSAKQTQTA  
KVTEASKLLGRAGLRAKVDLTGSSMTDLSSGGFTSGVATKGNELGLAEVANTI  
VKSSNLISLSKRNIEHLKGTILYSEGQNLVSNDFDELLRLVAADKROELQVFSGEV  
VRFNRSKDFQWHLNLRYPDRISKELTTPQRLKEDAVLVQDLVVLVQYTAELYQSLQ  
VLIRLEKDYEQKRESEANSKSGDLALATKELKAKVKVSKLKKSLNRSRGFEV  
MEKLDIVHFLLEIHNIFGGADQPSKKGAAYDKRLGPAFLALHYANILVQIDTLV  
ARASITSNARDSLYQSLPPGKILALRSKIKSFNDKELSVTQIKDEMERTILHLV  
AGNTTKAHGFGWGEWANTGDTFTSKPSGGDILRIETLYHASKKTEYIILQIITWL  
OHLVTKAKSDARGGPRLSISKSPDITNQQLISEPLSVPIVTDDEQKMLQEASKRKT  
PCVSKQDFDSEYSPARKCDPLSKSEYFRGVRSKSAVKRYSSGFFLLDPAIDKEK  
VLVDVDRVDVPRDYKALLKEGSLP"  
15981..16289  
/gene="At5g08660"  
/number=1  
16290..16432  
/gene="At5g08660"  
/number=1  
16433..16746  
/gene="At5g08660"  
/number=2  
16747..16832  
/gene="At5g08660"  
/number=2  
16833..16925  
/gene="At5g08660"  
/number=3  
16926..17019  
/gene="At5g08660"  
/number=3  
17020..17110  
/gene="At5g08660"  
/number=4  
17111..17304  
/gene="At5g08660"  
/number=4  
17305..17398  
/gene="At5g08660"  
/number=5  
17399..17477  
/gene="At5g08660"  
/number=5  
17478..17575  
/gene="At5g08660"  
/number=6  
17576..17766  
/gene="At5g08660"  
/number=6  
17767..17836  
/gene="At5g08660"  
/number=7  
17837..18027  
/gene="At5g08660"  
/number=7  
18028..18131  
/gene="At5g08660"  
/number=8  
18132..18230  
/gene="At5g08660"  
/number=8  
18231..18347



```

/gene="At5g08660"
/number=9
18348..18457
/gene="At5g08660"
/number=9
18458..18534
/gene="At5g08660"
/number=10
18535..18635
/gene="At5g08660"
/number=10
18636..18680
/gene="At5g08660"
/number=11
18681..18762
/gene="At5g08660"
/number=11
18763..19300
/gene="At5g08660"
/number=12
19850..22604
/gene="At5g08670"
complement(join(19850..19963,20204..20515,20610..20834,
20918..21076,21177..21280,21360..21477,21569..21657,
21738..21813,22131..22604))
/gene="At5g08670"
complement(join(19850..19963,20204..20515,20610..20834,
20918..21076,21177..21280,21360..21477,21569..21657,
21738..21813,22131..22604))
/gene="At5g08670"
/notes="strong similarity to H+-transporting ATP synthase,
mitochondrial, rubber tree, PIR:S20504
Contains ATP/GTP-binding site motif A (P-loop)
AA231-238:ATP synthase alpha and beta subunits signature
AA423-432"
/codon_start=1
/product="H+-transporting ATP synthase beta chain
(mitochondrial)-like protein"
/protein_id="CAC35872.1"
/db_xref="GI:13548325"
/db_xref="COA:P83483"
/db_xref="Swiss-Prot:P83483"
/translation="MASRRVLSLLRSSGRSAAKLNRPRLPSPARHAAPCSYL
LGRVATSPASAPAPAKDEGKTYDYGKAIGRVQVIGCAIVDFEQQE
GLPIITSLBQDHPTRLVLESHHGLQNVVRIAMDGTGLVGRKVLNTGAPITVP
VGRATLGRINNVLEPDERGEIKTEHLYPIHRDAPALVDLATGQRTLATGIKVDLL
APVQGGKIGLFGAGVGKTVLIMELINNVAKHGGFSPAGVGEERTRENDLIYREMI
ESGVIKIGESKSCALVYQWNPPEGARVGLTGLTVAEYFRDAEQDVLFDIN
IFRPTQANSVSALLGRIPSAVGQPTLASDLGALQERITTKKSGITSVQAIYVPAD
DLTDPATTFADLDATTVLSRQISLGIYPVDPDLDSTSRMLSPHILGEEHYNTARG
VQVLQNYKNLDIIATLGMDELSDDKLTVARARKIOLFSPFHVAEFTGAPGKY
VDLKENINSFGLLDKYDLSQSFFYWGVIDEVAKAKIASEAA"
complement(19850..19963)
/gene="At5g08670"
/number=1
complement(19964..20203)
/gene="At5g08670"
/number=1
complement(20204..20515)
/gene="At5g08670"
/number=2
complement(20516..20609)
/gene="At5g08670"
/number=2
complement(20610..20834)
/gene="At5g08670"
/number=3
complement(20835..20917)
/gene="At5g08670"
/number=3
complement(20918..21076)
/gene="At5g08670"
/number=4

```

```

intron complement(21077..21176)
/gene="At5g08670"
/number=4
exon complement(21177..21280)
/gene="At5g08670"
/number=5
intron complement(21281..21359)
/gene="At5g08670"
/number=5
exon complement(21360..21477)
/gene="At5g08670"
/number=6
intron complement(21478..21568)
/gene="At5g08670"
/number=6
exon complement(21569..21657)
/gene="At5g08670"

Query Match 9.9%; Score 62.4; DB 8; Length 82896;
Best Local Similarity 72.3%; Pred. No. 0.12;
Matches 81; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 160 CAGATATCGAAGAACAGGCGGTGCGGACCTGGAAACGCGAAGGTGCTGCTGCAC 219
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 77539 CAGATATCTAAGTAAGAGAGTCTCTGTAGGAANTGGAAACGGAAGATTGGATTGAC 77480
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 220 ATCCGCGAGTTCACGAGAGGACGCGACGACCTCCCGCGCGCAAGGTA 271
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 77479 ATTGGTGAGTTCTATGTCAAGGACGGAAGACATTGCTGCGCAAGGTA 77428
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
ATT5E8/c
LOCUS 103534 bp DNA linear PLN 29-AUG-2000
DEFINITION Arabidopsis thaliana DNA chromosome 5, BAC clone T5E8 [ESSA
project].
ACCESSION AL391712
VERSION AL391712.1 GI:9955507
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
1 (bases 1 to 103534)
Bevan,M., Murphy,G., Ridley,P., Hudson,S., Bancroft,I., Mewes,H.W.,
Rudd,S., Lemcke,K. and Mayer,K.F.X.
Unpublished
JOURNAL
2 (bases 1 to 103534)
EU Arabidopsis sequencing,project.
AUTHORS
TITLE
JOURNAL
Submitted (28-AUG-2000) MIPS, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
E-mail: michael.bevan@bbsrc.ac.uk
COMMENT
Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.
FEATURES
Location/Qualifiers
source
1..103534
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="5"
3395..6144
/gene="T5E8 10"
complement(join(3395..3451,3525..3633,4011..4306,
4412..4685,4910..5226,5275..5397,5442..5567,5635..6144))
/gene="T5E8 10"
complement(join(3395..3451,3525..3633,4011..4306,
complement(

```

```

4412..4685,4910..5226,5275..5397,5442..5567,5635..6144))
/genes="T5E8_10"
/notes="similarity to GCF2 fusion protein - Homo sapiens,
EMBL:AB026911"
/codon_start=1
/product="putative protein"
/protein_id="CAC05447.1"
/db_xref="GI:9955508"
/db_xref="T:EMBL:Q9FY92"
/translation="MCTKOPRYRRRCNDEIDGEDATAAKPSSSDLYPAKPSDDLY
PKRRLKLPATPKNDGDAKGRHAKSSWLLDLSRGDEFTYQOMQORLKEHQDERPM
ELVFNGBRSGDGLSTVFVCPMPQTADDKMAALDRICNAYDERVNRMPQEQPM
YAGRPELAKESPDENLTGIAAECSAPSVMYKASLQISDFKSVFRNPMQGI CVAF
VCVSGAFISIKLLSIFFLSQNGYLITAIEDQMKVDGYSLIVEGDSSTDDSDCETS
AYEERADSLQADKIFSDASVYVYSELRSVKIFKRGARHPSAPRAAYTSLTVPSMY
SPYLRLLELRMDPLHQDVDFSDMWHGLAFHSRIYVCGSTPVCNTNPFVSELKYAVAP
ILHRLRVQCDLILSTRETRNVNVAATSLVARYVFPSEALAELSLAIHARLVRAI AIS
VTPWDQKVPNAPQVAAVAFGTSVRLMRNICMVKOMVLELPVLEKLSALLFGKV
LPHVRSIASESNIHDAVTKTERIVASLSGVWTGSPVTRTHSHLLQLPLVDCTLTGLRIL
EKVKCLIGTGRYGNHMSRPQEDVDLQLTDPFVSRIKVI"
/complement (3195..3451)
/genes="T5E8_10"
/number=1
/complement (3452..3524)
/genes="T5E8_10"
/number=1
/complement (3525..3633)
/genes="T5E8_10"
/number=2
/complement (3634..4010)
/genes="T5E8_10"
/number=2
/complement (4011..4306)
/genes="T5E8_10"
/number=3
/complement (4307..4411)
/genes="T5E8_10"
/number=3
/complement (4412..4685)
/genes="T5E8_10"
/number=4
/complement (4686..4909)
/genes="T5E8_10"
/number=4
/complement (4910..5226)
/genes="T5E8_10"
/number=5
/complement (5227..5274)
/genes="T5E8_10"
/number=5
/complement (5275..5397)
/genes="T5E8_10"
/number=6
/complement (5398..5441)
/genes="T5E8_10"
/number=6
/complement (5442..5567)
/genes="T5E8_10"
/number=7
/complement (5568..5634)
/genes="T5E8_10"
/number=7
/complement (5635..6144)
/genes="T5E8_10"
/number=8
7652..9648
/genes="T5E8_20"
join(7652..7793,7900..8133,8215..8308,8403..8617,
8698..8837,8992..9648)
/notes="Contains Prokaryotic membrane lipoprotein lipid
attachment site AA454-464"
/codon_start=1
/product="amino acid transport protein AAP2"
/protein_id="CAC05448.1"
/db_xref="GI:9955509"
/db_xref="GOA:Q38967"
/db_xref="T:EMBL:Q38967"
/translation="MGETAANNRRHHHHHHHGHQVFDVASHDFVPPQAPKCEDDDGRL
KRTGTWTASAHITTAIVAGSGVLAWIAQIAGWAGPAVMLLFLSVLTYSSTLLSDC
YRTDAVSGKRNITVMDAVNSYILGGFKFKICGLIOVLNFIAGIAGTYIAASIMMAIK
RNCDFHSGGKDPCHMSSNPYIMVFGVAEILLQVDFDQIWIWISIVAAVSVLYEIQD
GLALGIVOAVANGVFKGSLTGISICTVCTOKIMRTFOALGDIAPAYSYSVVLIEIQD
TVRSPASKTNKKATKISIAVTTTFYMLCGSMGYAAFGDAAPGNLLTGFQYNPFWL
LDIAANAIVHLVGAIVQVPAQFIPIEKSAERYPDNDFLSKEFEIRIPGFKSPYKV
NIFRVYRSGFVVTTVISMLPFNVVGIILGALGFELTVPYFPVEMYIKORKVEKM
STRWVCLQWLSVACLIVISVAVGSGIAGVMLDKVYKPKKSTY"
/genes="T5E8_20"
/number=1
7794..7899
/genes="T5E8_20"
/number=1
7900..8133
/genes="T5E8_20"
/number=2
8134..8214
/genes="T5E8_20"
/number=2
8215..8308
/genes="T5E8_20"
/number=3
8309..8402
/genes="T5E8_20"
/number=3
8403..8617
/genes="T5E8_20"
/number=4
8618..8697
/genes="T5E8_20"
/number=4
8698..8837
/genes="T5E8_20"
/number=5
8838..8991
/genes="T5E8_20"
/number=5
8992..9648
/genes="T5E8_20"
/number=6
11840..14398
/genes="T5E8_30"
join(11840..11995,12106..12214,12604..12841,12924..12989,
13090..13135,13245..13405,13482..13617,13691..13855,
13929..14056,14129..14228,14348..14398)
/genes="T5E8_30"
/notes="similarity to siruin type 4 (SIRT4) - Homo
sapiens, EMBL:AF083109"
/codon_start=1
/product="SIR2-family protein"
/protein_id="CAC05449.1"
/db_xref="GI:9955510"
/db_xref="GOA:Q9FY91"
/db_xref="T:EMBL:Q9FY91"
/translation="MKLHQDKNRRESHERIFORELTRSSSRERDSEGEIAKERELR
PQRENKIDTDFCHLVFNNNKKFKIKRKTETISFHPPLTTSQLFPPSRSMVYRPLQS
GSLNMLFKGCRFRVTCRVISIPGSGLSNESKAPRFLDRDKIVPDADPPNMEDIHK
LYRLFEQSRRLTILTAGVSTECGIPDYRSPNGAYSSGFKPITHGFTTRSSARRRYV
ARSYAGWRRRTAAQPGPAHTALASLEKAGRNFMITQNVDRLUHRAGSDPFLHGTVY
TVMCLECGFSFPDRILFQDLKAINPKASWAEAIIESIDHGDGPGSEKSGFMKQPDGDI
IDEKVEEGFPIPVCEKCKGVLPDVIFFGDNIPKERATQAMEVAKQSDAFLVLGSSL
MTMSAFRLCAAHEAGAMTAIWNIGETRADDIVPLKINARVGEILHRLVLDVLSLSPA
L"
11840..11995
/genes="T5E8_30"
/number=1

```

intron 11996..12105  
/gene="TSE8\_30"  
/number=1  
exon 12106..12214  
/gene="TSE8\_30"  
/number=2  
intron 12215..12603  
/gene="TSE8\_30"  
/number=2  
exon 12604..12841  
/gene="TSE8\_30"  
/number=3  
intron 12842..12923  
/gene="TSE8\_30"

Query Match 9.9%; Score 62.4; DB 8; Length 103534;  
Best Local Similarity 72.3%; Pred. No. 0.12;  
Matches 81; Conservative 0; Mismatches 31; Indels 0; Gaps 0;  
Qy 160 CAGATATCGAAGAACACGAGGGTGGCGGTGCGGACCTGGAAACGGCAAGGTCGTGTCGAC 219  
Db 16869 CAGATATCTAAGAATAGGAGAGTCTCTGAAGGAATTGGAACGGGAAGATTGGATTGAC 16810  
Qy 220 ATCGCGAGTTCTACGAGAGGACGGCAAGACCCTCCCGCCGCCAAGGTA 271  
Db 16809 ATTGCTGAGTTCTATGTCAAGGACGGAAAGACTTTGCCTGGCAAGAAAGGTA 16758

Search completed: January 11, 2005, 21:05:45  
Job time : 3034 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 11, 2005, 12:05:19 ; Search time 441 Seconds  
(without alignments)  
7522.980 Million cell updates/sec

Title: US-10-629-953-3  
Perfect score: 632  
Sequence: 1 ctctccagactcccaag.....taaaaaaaaaaaaaaaaaaaaaa 632

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_23Sep04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	632	100.0	632	3	Aaz50412 Rice tran
2	220.4	34.9	460	3	Aaz50414 Wheat tra
3	210.6	33.3	649	3	Aaz50411 Corn tran
4	182.2	28.8	289	10	Abx82999 Corn ear-
5	149.6	23.7	263	10	Abx83152 Corn ear-
6	146.6	23.2	310	6	Abx83152 Corn ear-
7	145.6	23.0	286	10	Abx88456 Corn ear-
8	109.4	17.3	672	3	Aaz50413 Soybean t
9	96	15.2	471	3	Aac38126 Arabidops
10	78.6	12.4	516	3	Aaz50416 Vernonia
11	76.4	12.1	939	3	Aaz50418 Corn tran
12	76	12.0	1089	3	Aaz50417 Corn tran
13	74	11.7	1063	3	Aac32760 Arabidops
14	68.2	10.8	913	8	Acf03526 Brassica
15	66.2	10.5	740	3	Aaz50419 Soybean t
16	61.6	9.7	498	3	Aaz50415 Marigold
17	61	9.7	496	10	Ades1749 Arabidops
18	61	9.7	757	3	Aac50198 Arabidops
19	61	9.7	757	3	Aac34846 Arabidops
20	55.2	8.7	2000	8	Ada71938 Rice gene
21	52.4	8.3	1074	3	Aaa76284 Maize glu

## ALIGNMENTS

RESULT 1  
AAZ50412

ID AAZ50412 standard; cDNA; 632 BP.

XX AC AAZ50412;

XX AC (first entry)

XX 18-MAY-2000 (first entry)

XX Rice transcription coactivator PC4(P15) type 1 cDNA.

XX Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 1;  
XX rice; signal mediator; activator; Gal4/VP16; transgenic plant;  
XX general transcription factor; GTF; transcription initiation complex;  
XX immunological screening; detection; marker; clone rri.pk0003.a12; ss.

XX Oryza sativa.

XX Key Location/Qualifiers

XX CDS 46..351

XX FT /\*tag= a

XX FT product= "Rice PC4(P15) type 1 transcription

XX FT coactivator"

XX FT /note= "Derived from clone rri.pk0003.a12"

XX WO200005377-A2.

XX 03-FEB-2000.

XX 21-JUL-1999; 99WO-US016479.

XX 22-JUL-1998; 98US-0093687P.

XX (DUPO ) DU PONT DE NEMOURS & CO E I.

XX Cahoon RE, Caimi PG, Odell JT, Sakai H, Zhu Q;

XX WPI; 2000-182701/16.

XX P-PSDB; AAY44879.

XX Novel PC4 transcriptional coactivator polynucleotides and polypeptides

XX used to alter the level of PC4(P15) type 1 and PC4(P15) type 2

XX polynucleotides and polypeptides.

XX Claim 9; Page 33; 43pp; English.

Adc72476 DNA Seq I  
Adc72595 DNA Seq I  
Add17205 DNA (Seq)  
Adk58353 Plant DNA  
Ab139958 Synthetic  
Adm73763 HIV-1 pol  
Adj44561 Plant cDN  
Abq81847 Bifidobac  
Acf35993 Erythrocy  
Acf35992 Erythrocy  
Aaa51610 HIV synth  
Aaa51626 HIV codon  
Aal44553 HIV-1 p55  
Aal44549 HIV-1 p55  
Ab139954 Synthetic  
Adm73759 HIV-1 pol  
Adn36395 Human gen  
Adn36401 Human gen  
Adm7600 Polynucle  
Abd03958 Pseudomon  
Aaa51578 Pseudomon  
Aaa19447 Prokaryot  
Adj39777 Plant cDN  
Abd03877 Pseudomon

22 49.2 7.8 458 10 ADC72476  
23 49.2 7.8 458 10 ADC72595  
24 49.2 7.8 458 10 ADD17205  
25 49.2 7.8 458 10 ADK58353  
26 48 7.6 1509 6 ABL39958  
27 48 7.6 1509 12 ADM73763  
28 47.6 7.5 537 12 ADJ44561  
29 47.4 7.5 349980 6 ABQ81847  
30 46.8 7.4 1848 9 ACF35993  
31 46.8 7.4 1848 9 ACF35992  
32 46.4 7.3 1509 3 AAA51610  
33 46.4 7.3 1509 3 AAA51626  
34 46.4 7.3 1509 6 AAL44553  
35 46.4 7.3 1509 6 AAL44549  
36 46.4 7.3 1509 6 ABL39954  
37 46.4 7.3 1509 12 ADM73759  
38 46.4 7.3 1545 12 ADN36395  
39 46.4 7.3 8186 12 ADN36401  
40 46 7.3 1795 12 ADM7600  
41 45.8 7.2 954 11 ABD03958  
42 45.8 7.2 1416 4 AAS51578  
43 45.8 7.2 1416 8 AAC19447  
44 45.8 7.2 1458 12 ADJ39777  
45 45.8 7.2 1953 11 ABD03877

XX The present sequence is the cDNA encoding rice Positive Cofactor 4 (PC4)  
CC transcription coactivator, designated as PC4(p15) type 1. It is isolated  
CC from clone rr1.pK0003.a12, obtained from rr1 cDNA library prepared using  
CC rice root two week old developing seedlings. PC4 molecules functions as a  
CC signal mediator between activators like Gal4/VP16 and general  
CC transcription factors (GTFs) in a transcription initiation complex. The  
CC expression levels of PC4 can be manipulated and the functional properties  
CC of specific transcriptional activators can be modulated. The PC4 DNA can  
CC be used to create transgenic plants with altered PC4 levels, that would  
CC affect the level of transcription of specific genes in the plant. It is  
CC also used for immunological screening of cDNA libraries and to raise  
CC specific antibodies for detection. The DNA is used as probes and primers,  
CC for genetic and physical mapping of genes and as markers for traits  
CC linked to those genes  
XX  
SQ Sequence 632 BP; 165 A; 139 C; 187 G; 141 T; 0 U; 0 Other;

DT	18-MAY-2000	(first entry)
XX	Wheat transcription coactivator PC4(P15) type 1 cDNA.	
XX	Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 1;	
KW	wheat; signal mediator; activator; Gal4/VP16; transgenic plant;	
KW	general transcription factor; GTF; transcription initiation complex;	
KW	immunological screening; detection; marker; clone wdk2c.pk015.g20; ss.	
XX	Triticum aestivum.	
OS		
XX	Key Location/Qualifiers	
PH	3..293	
FT	CDS	
FT	/tag= a	
FT	/product= "Wheat PC4(P15) type 1 transcription	
FT	coactivator"	
FT	/note= "Derived from clone wdk2c.pk015.g20"	
XX	WO200005377-A2.	
PN		
XX	03-FEB-2000.	
PD		
XX		
PF	21-JUL-1999; 99WO-US016479.	
XX		
PR	22-JUL-1998; 98US-0093687P.	
XX		
PA	(DUPO ) DU PONT DE NEMOURS & CO E I.	
XX		
PI	Cahoon RE, Caimi PG, Odell JT, Sakai H, Zhu Q;	
XX		
DR	WPI; 2000-182701/16.	
DR	P-PSDB; AAY44881.	
XX		
PT	Novel PC4 transcriptional coactivator polynucleotides and polypeptides	
PT	used to alter the level of PC4(P15) type 1 and PC4(P15) type 2	
PT	polynucleotides and polypeptides.	
XX		
XX	Claim 3; Page 35; 43pp; English.	
XX		
CC	The present sequence is the cDNA encoding wheat Positive Cofactor 4 (PC4)	
CC	transcription coactivator, designated as PC4(P15) type 1. It is isolated	
CC	from clone wdk2c.pk015.g20, obtained from wdk2c cDNA library, prepared	
CC	from wheat developing kernels, 7 days after anthesis. PC4 molecules	
CC	function as a signal mediator between activators like Gal4/VP16 and	
CC	general transcription factors (GTFs) in a transcription initiation	
CC	complex. The expression levels of PC4 can be manipulated and the	
CC	functional properties of specific transcriptional activators can be	
CC	modulated. The PC4 DNA can be used to create transgenic plants with	
CC	altered PC4 levels, that would affect the level of transcription of	
CC	specific genes in the plant. It is also used for immunological screening	
CC	of cDNA libraries and to raise specific antibodies for detection. The DNA	
CC	is used as probes and primers, for genetic and physical mapping of genes	
CC	and as markers for traits linked to those genes	
XX		
SQ	Sequence 460 BP; 117 A; 110 C; 139 G; 87 T; 0 U; 7 Other;	
	Query Match 34.9%; Score 220.4; DB 3; Length 460;	
	Best Local Similarity 79.3%; Pred. No. 2.7e-47;	
	Matches 260; Conservative 0; Mismatches 68; Indels 0; Gaps 0;	
QY	72 CGCGCGCGCGCGAGCCGGCGGCGAAGCCCGTGCCTGCCGCGGGAACGAGGGCCTCCGA 131	
Db	14 CGCGCGCGCGCGACGCCCGCGCAAGCCAGCCAGCGCGCGCGGAGGACGGTCCCTCCGA 73	
QY	132 GAGCGCGGACGACGATATCGTCGTGCCCCAGATATCGAAGMACAGAGGGTGCGGTGCG 191	
Db	74 GGAACCGACGACGGCATCGTCGTTCGCAGATATCGAAGAACAAAGAGGGTGCCGTGAG 133	
QY	192 GACTCTGMAAGCGCAAGTCGTCTGTCACATCCCGAGTTCTACGAGAAGGACGGCAAGAC 251	
Db	134 GAACCTGMAAGCGGAAGGTCATGCTGCAATGCGCGAGTTCTACGAAAAGGACGGCAAGAG 193	
QY	252 CCTCCCCCGCGCAAAGGTATACAGCTCCCATTGATCGATGGAAGATCTGAGGACAA 311	

```

||||| 194 CCTCCGACCGCAAGATATATCGCTCTCAATGGATCAGTGGAAATATCTGAGGGCAA 253
||||| 312 TATCAAGCTATAGATGAGGCGCATCAAGGAGAAATCGCTGATCGGAGGCCATTCTCTTGTG 371
||||| 254 CATCGAGCTATAGACGAGGCCATCAAGGAGAACACTTGATCAGAAAAGCGGTTAAGGAN 313
||||| 372 ATGCAAGTAGACTAAGCTAGCTGTCTC 399
||||| 314 ATNTTACTGGGCAAAAGTGTGCTCATTC 341

```

## RESULT 3

AAZ50411

ID AAZ50411 standard; cDNA; 649 BP.

XX AC

AAZ50411;

XX AC

DT 18-MAY-2000 (first entry)

XX DT

Corn transcription coactivator PC4 (P15) type 1 cDNA.

XX DE

Positive cofactor 4; PC4; transcription coactivator; PC4 (P15) type 1;  
 corn; signal mediator; activator; Gal4/VP16; transgenic plant;  
 general transcription factor; GFP; transcription initiation complex;  
 immunological screening; detection; marker; clone cca.pk0020.d2; ss.

XX KW

XX OS

XX ZE

Key Location/Qualifiers

XX PH

XX CDS

FT 33..344

FT /\*tag= a

FT /product= "Corn PC4 (P15) type 1 transcription

FT coactivator"

FT /note= "Derived from clone cca.pk0020.d2"

XX FT

XX PN

XX WO200005377-A2.

XX PD

XX 03-FEB-2000.

XX PF 21-JUL-1999; 99WO-US016479.

XX PR 22-JUL-1998; 98US-0093687P.

XX PA (DUPO ) DU PONT DE NEMOURS &amp; CO E I.

XX PI Cahoon RE, Caimi PG, Odell JT, Sakai H, Zhu Q;

XX XX

XX WPI; 2000-182701/16.

XX DR P-PSDB; AAY44878.

XX XX

Novel PC4 transcriptional coactivator polynucleotides and polypeptides  
 used to alter the level of PC4 (P15) type 1 and PC4 (P15) type 2  
 polynucleotides and polypeptides.

XX PT

XX PS

Claim 3; Page 32; 43pp; English.

XX XX

The present sequence is the cDNA encoding corn Positive Cofactor 4 (PC4)  
 transcription coactivator, designated as PC4 (P15) type 1. It is isolated  
 from the corn callus type II tissue (undifferentiated) from the cDNA  
 clone cca.pk0020.d2. PC4 molecules functions as a signal mediator between  
 activators like Gal4/VP16 and general transcription factors (GTFs) in a  
 transcription initiation complex. The expression levels of PC4 can be  
 manipulated and the functional properties of specific transcriptional  
 activators can be modulated. The PC4 DNA can be used to create transgenic  
 plants with altered PC4 levels, that would affect the level of  
 transcription of specific genes in the plant. It is also used for  
 immunological screening of cDNA libraries and to raise specific  
 antibodies for detection. The DNA is used as probes and primers, for  
 genetic and physical mapping of genes and as markers for traits linked to  
 those genes

XX CC

Sequence 649 BP; 167 A; 147 C; 182 G; 153 T; 0 U; 0 Other;

XX SQ

```

Query Match 33.3%; Score 210.6; DB 3; Length 649;
Best Local Similarity 77.2%; Pred. No. 1.1e-44;
Matches 271; Conservative 0; Mismatches 74; Indels 6; Gaps 1;

Qy 15 CACAACTCCAGAGGAGAAAGTGTGAGAAAGTGTGCGGGAAGGGAACAAGCGGTTCCG 74
Db 2 CACGAGGTCCAGTCCGAGTCTCAGCAAGGAAAGATGTGGGGGAAAGAAAGCGTTTCGG 61
Qy 75 CGGCGGCGGC-----GAGCGGCGGCAAGCGCGCTGCGCGGGGACGACGGGCCCTC 128
Db 62 CGGTGGCGGACGCGGTGAGCCCGCGGCAAGCGCGGCGCGGAGGACGATGGCCCTC 121
Qy 129 CGAGAGCGCCGACGACGATATCGTCTGCGCCGAGATATCGAAGAACAGAGGGGTGGCGGT 188
Db 122 CGAATCGCTGAAGACGGTACCGTCTGCTGCGAGTATATCAGCGAGTTCTACTTTCAAGGACGCAA 241
Qy 189 GCGGACCTGGAACGGCAAGGTCTGCTGCGACATCCGCGAGTTCTACGAGAAGGACGGCAA 248
Db 182 TAGGAGCTGGAAGGCGAGGGTCTTCTGCTGAGTTCACGCGAGTTCTACTTTCAAGGACGCAA 241
Qy 249 GACCTCCCGCGCGCAAGGTATACAGCTCCCAATGGATCAGTGGGAAGATCTAGAGGA 308
Db 242 GACTCTCCCGCGCAAGGTATATCAGTCCCAATTAGATCAGTGGGAAGATCTAGAGGA 301
Qy 309 CAATATCAAAAGCTATAGATGAGGCGCATCAAGGAGAAATGCGTGATCGGAGCC 359
Db 302 CAACATCAAGCCATAAATGAGGCAATCGAGGAATAATACATGATGGAGCC 352

```

## RESULT 4

ABX82999

ID ABX82999 standard; cDNA; 289 BP.

XX AC

ABX82999;

XX DT

24-APR-2003 (first entry)

XX DE

Corn ear-derived polynucleotide (cpd) #1459.

XX XX

Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; SATMON023;  
 structural gene; functional gene; regulatory gene;  
 corn ear-specific profile; gene transcription; gene expression;  
 hybrid plant; desirable trait expression; plant breeding program;  
 inheritance; desired characteristic; growth; development;  
 disease resistance; environmental adaptability; quality; yield;  
 multigene trait; plant; gene; ss.

XX OS

XX ZE

Zea mays.

XX US6476212-B1.

XX PN

XX XX

XX PD

XX 05-NOV-2002.

XX XX

XX PF

XX 14-MAY-1999; 99US-00313294.

XX XX

XX PR

XX 26-MAY-1998; 98US-0086722P.

XX XX

XX PA

XX (INCY-) INCYTE GENOMICS INC.

XX XX

XX PI

XX Lalgudi RV, Ito LY, Sherman BK;

XX XX

XX WPI; 2003-208840/20.

XX XX

XX PT

XX Novel purified corn-ear derived polynucleotide useful as hybridization

XX PT

XX probe for detecting polynucleotide in sample, and for identifying,

XX PT

XX evaluating, and altering desired characteristics associated with growth,

XX XX

XX development.

XX XX

XX PS

XX Example; SEQ ID NO 1459; 390pp; English.

XX XX

XX CC

XX The present invention relates to the isolation of corn ear-derived

XX CC

XX polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022

CC and SATMON023. Some of the cpds uniquely identify structural, functional, CC and regulatory genes of corn ear. The polynucleotides sequences are CC useful for detecting cpds in a sample, for producing a corn ear-specific CC profile of gene transcription, for detecting altered gene expression in CC inbred or hybrid plants, and for screening several molecules for specific CC binding to the polynucleotide. The cpds are useful to identify, isolate, CC or extend identical or related corn-ear nucleic acid sequences from DNA CC libraries, and in nucleic acid amplification or hybridisation techniques CC to follow the expression of desirable traits through plant breeding CC programs. Preferably, the cpds are used to identify, evaluate, alter, or CC follow the inheritance of desired characteristics associated with growth CC and development, disease resistance, environmental adaptability, quality, CC and yield of corn. The cpds are also useful as molecular markers for CC studying inheritance and multigene traits in a plant breeding program. CC The cpds are useful for producing purified corn-ear polypeptides by CC recombinant techniques. They are also useful in diagnostic assays to CC detect or confirm conditions or diseases associated with abnormal levels CC of cdp expression. ABX81541-ABX89140 represent corn ear-derived CC polynucleotides (cpds) of the invention. Note: The sequence data for this CC patent did not form part of the printed specification, but was obtained CC in electronic format directly from the USPTO web site at CC seqdata.uspto.gov/psipdbIDEntry.html

XX SQ Sequence 289 BP; 79 A; 71 C; 92 G; 47 T; 0 U; 0 Other;  
Query Match 28.8%; Score 182.2; DB 10; Length 289;  
Best Local Similarity 79.8%; Pred. No. 2e-37;  
Matches 230; Conservative 0; Mismatches 53; Indels 6; Gaps 1;  
QY 41 GAAGATGTGGCGGAGGGAACAAGCGGTTCGGCGCGCGCGC-----GAGCCGCGG 94  
DB 1 GGAAGATGTGGGGAAGGAAGACGCTTTCGGCGGTGAGCCGCGG 60  
QY 95 CCAAGCGCGTCCGCGGAGGACGAGCGGCTTCGAGAGCGCGGACGATATCGTCG 154  
DB 61 CCAAGCGCGGCGGAGGAGCGATGGCCCTCCGAATCCGCTGAAGACGGTACCGTCG 120  
QY 155 TCGCCAGATATCGAAGACAGGAGGTGGCGGTTCGGACCTCGAAGCGCAAGTCTGTCG 214  
DB 121 TAGCCGAGATATCGAAGACAGAGGTGTCCGTTAGGAGCTGGAAGCGAGGTCTTCG 180  
QY 215 TCACATCCGCGAGTTCTACGAGAAGGACGCGAAGACCCCTCCCGCGCGCAAGGTATAC 274  
DB 181 TCAGCTTACGCGAGTTCTACTTCAAGGACGCAAGACTCTCCCGACCCGCAAGGTATAT 240  
QY 275 AGCTCCCAATGATCAGTGAAGACTAGGACACATATCAAGCTAT 323  
DB 241 CACTCCAATTAGATCAGTGGAGATATTGAAGGACACATCAAGGCCAT 289

RESULT 5  
ABX83152  
ID ABX83152 standard; cDNA; 263 BP.  
XX AC ABX83152;  
XX AC  
DT 24-APR-2003 (first entry)  
XX Corn ear-derived polynucleotide (cpd) #1612.  
DE  
XX Corn ear-derived polynucleotide; cdp; cDNA library; SATMON023;  
KW structural gene; functional gene; regulatory gene;  
KW corn ear-specific profile; gene transcription; gene expression;  
KW hybrid plant; desirable trait expression; plant breeding program;  
KW inheritance; desired characteristic; growth; development;  
KW disease resistance; environmental adaptability; quality; yield;  
KW multigene trait; plant; gene; ss.  
XX  
OS Zea mays.  
XX  
FN US6476212-B1.  
XX  
XX 05-NOV-2002.

XX 14-MAY-1999; 99US-00313294.  
XX 26-MAY-1998; 98US-0086722P.  
XX (INCY-) INCYTE GENOMICS INC.  
XX Lalgudi RV, Ito LY, Sherman BK;  
XX WPI; 2003-208840/20.  
DR Novel purified corn-ear derived polynucleotide useful as hybridization  
XX probe for detecting polynucleotide in sample, and for identifying,  
PT evaluating, and altering desired characteristics associated with growth,  
PT development.  
XX Example; SEQ ID NO 1612; 390pp; English.  
PS The present invention relates to the isolation of corn ear-derived  
XX polynucleotides (cpds) from the corn (Zea mays) cDNA libraries SATMON022  
CC and SATMON023. Some of the cpds uniquely identify structural, functional,  
CC and regulatory genes of corn ear. The polynucleotides sequences are  
CC useful for detecting cpds in a sample, for producing a corn ear-specific  
CC profile of gene transcription, for detecting altered gene expression in  
CC inbred or hybrid plants, and for screening several molecules for specific  
CC binding to the polynucleotide. The cpds are useful to identify, isolate,  
CC or extend identical or related corn-ear nucleic acid sequences from DNA  
CC libraries, and in nucleic acid amplification or hybridisation techniques  
CC to follow the expression of desirable traits through plant breeding  
CC programs. Preferably, the cpds are used to identify, evaluate, alter, or  
CC follow the inheritance of desired characteristics associated with growth  
CC and development, disease resistance, environmental adaptability, quality,  
CC and yield of corn. The cpds are also useful as molecular markers for  
CC studying inheritance and multigene traits in a plant breeding program.  
CC The cpds are useful for producing purified corn-ear polypeptides by  
CC recombinant techniques. They are also useful in diagnostic assays to  
CC detect or confirm conditions or diseases associated with abnormal levels  
CC of cdp expression. ABX81541-ABX89140 represent corn ear-derived  
CC polynucleotides (cpds) of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from the USPTO web site at  
CC seqdata.uspto.gov/psipdbIDEntry.html  
XX SQ Sequence 263 BP; 67 A; 65 C; 88 G; 41 T; 0 U; 2 Other;

Query Match 23.7%; Score 149.6; DB 10; Length 263;  
Best Local Similarity 78.0%; Pred. No. 6e-29;  
Matches 206; Conservative 0; Mismatches 51; Indels 7; Gaps 2;  
QY 41 GAAGATGTGGCGGAGGGAACAAGCGGTTCGGCGCGCGCGC-----GAGCCGCGG 94  
DB 1 GGAAGATGTGGGGAAGGAAGACGCTTTCGGCGGTGAGCCGCGG 60  
QY 95 CCAAGCGCGTCCGCGGAGGACGAGCGGCTTCGAGAGCGCGGACGATATCGTCG 154  
DB 61 CCAAGCGCGGCGGAGGAGCGATGGCCCTCCGAATCCGCTGAAGACGGTACCGTCG 120  
QY 155 TCGCCAGATATCGAAGACAGGAGGTGGCGGTTCGGACCTCGAAGCGCAAGTCTGTCG 214  
DB 121 TAGCCGAGATATCGAAGACAGAGGTGTCCGTTAGGAGCTGGAAGCGAGGTCTTCG 180  
QY 215 TCACATCCGCGAGTTCTACGAGAAGGACGCGCAAGACCCCTCCCGCGCGCAAGGTATAC 274  
DB 181 NCGACTTACGCGAGTTCTACTTCAAGGACGCGCAAGACTCTCCCGACCCGCAAGGTATA- 239  
QY 275 AGCTCCCAATGATCAGTGAAGACTAGGACACATATCAAGCTAT 298  
DB 240 TCATCCAATTAGATCAGTGAAGACTAGGACACATATCAAGGCCAT 263

RESULT 6  
ABL70693  
ID ABL70693 standard; cDNA; 310 BP.





CC The cpds are useful for producing purified corn-ear polypeptides by  
 CC recombinant techniques. They are also useful in diagnostic assays to  
 CC detect or confirm conditions or diseases associated with abnormal levels  
 CC of cdp expression. ABX81541-ABX89140 represent corn ear-derived  
 CC polynucleotides (cpds) of the invention. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from the USPTO web site at  
 CC seqdata.uspto.gov/psipdIDEntry.html  
 XX  
 SQ Sequence 286 BP; 69 A; 77 C; 92 G; 47 T; 0 U; 1 Other;  
 Query Match 23.0%; Score 145.6; DB 10; Length 286;  
 Best Local Similarity 78.4%; Pred. No. 6.9e-28;  
 Matches 189; Conservative 0; Mismatches 45; Indels 7; Gaps 1;  
 QY 40 AGAAGATGTGGCGGGAAGGGAACAGCGTTTC-----GGCGGCGGCGGCGCGGC 92  
 DB 44 AGGAAGATGTGGGGAAGGGAAGGGAAGCGTTTCGGCGGTGGCGGCGGCGGCGGC 103  
 QY 93 GGCCAAAGCGCGTGGCGCGCGCGGACGACGGCGCTCCGAGAGCGCGGACGATATCGT 152  
 DB 104 GGCCAAAGCGCGCGCGGAGCGATGGCCCTCCGATCCGCTGAGACGGTACCGT 163  
 QY 153 CGTGGCCGAGATATCGAAGAACAGGAGGCGTGGCGGCGGACCTCGAAACGGCAAGTCT 212  
 DB 164 CGTAGCGGAGATATCGAAGAACAGGAGGCGTGGCGGCGGAGCTGGAAAGCGGAGTCT 223  
 QY 213 CGTGGACATCCGCGAGTTCTACGAGAGGAGGCGGAGCGGCGGCGGCGGCGGAGGAT 272  
 DB 224 CGTGGAGTACGCGGAGTTCTACGAGAGGAGGCGGAGCGGCGGCGGCGGCGGAGGAT 283  
 QY 273 A 273  
 DB 284 A 284  
 RESULT 8  
 AAZ50413  
 ID AAZ50413 standard; cDNA; 672 BP.  
 XX  
 AC AAZ50413;  
 XX  
 DT 18-MAY-2000 (first entry)  
 XX  
 DE Soybean transcription coactivator PC4(P15) type 1 cDNA.  
 XX  
 KW Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 1;  
 KW soybean; signal mediator; activator; Gal4/VP16; transgenic plant;  
 KW general transcription factor; GTF; transcription initiation complex;  
 KW immunological screening; detection; marker; clone sf11.pk0008.a4; ss.  
 XX  
 OS Glycine max.  
 XX  
 FH Location/Qualifiers  
 CDS 37..354  
 FT /\*tag= a  
 FT /product= "Soybean PC4(P15) type I transcription  
 FT coactivator"  
 FT /note= "Derived from clone sf11.pk0008.a4"  
 XX  
 XX WO200005377-A2.  
 XX  
 XX 03-FEB-2000.  
 XX  
 PF 21-JUL-1999; 99WO-US016479.  
 XX  
 XX 22-JUL-1998; 98US-0093687P.  
 XX  
 XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
 PA  
 XX Cahoon RE, Cai MI, PG, Odell JT, Sakai H, Zhu Q;  
 PI  
 XX WPI; 2000-182701/16.  
 DR

DR P-PSDB; AAY44880.  
 XX  
 XX Novel PC4 transcriptional coactivator polynucleotides and polypeptides  
 PT used to alter the level of PC4(P15) type 1 and PC4(P15) type 2  
 PT polynucleotides and polypeptides.  
 XX  
 PS Claim 3; Page 33-34; 43pp; English.  
 XX  
 CC The present sequence is the cDNA encoding soybean Positive Cofactor 4  
 CC (PC4) transcription coactivator, designated as PC4(P15) type 1. It is  
 CC isolated from clone sf11.pk0008.a4, obtained from sf11 cDNA library,  
 CC prepared from soybean immature flowers. PC4 molecules functions as a  
 CC signal mediator between activators like Gal4/VP16 and general  
 CC transcription factors (GTFs) in a transcription initiation complex. The  
 CC expression levels of PC4 can be manipulated and the functional properties  
 CC of specific transcriptional activators can be modulated. The PC4 DNA can  
 CC be used to create transgenic plants with altered PC4 levels, that would  
 CC affect the level of transcription of specific genes in the plant. It is  
 CC also used for immunological screening of cDNA libraries and to raise  
 CC specific antibodies for detection. The DNA is used as probes and primers,  
 CC for genetic and physical mapping of genes and as markers for traits  
 CC linked to those genes  
 XX  
 SQ Sequence 672 BP; 218 A; 113 C; 154 G; 187 T; 0 U; 0 Other;  
 Query Match 17.3%; Score 109.4; DB 3; Length 672;  
 Best Local Similarity 66.1%; Pred. No. 2.5e-18;  
 Matches 158; Conservative 0; Mismatches 81; Indels 0; Gaps 0;  
 QY 127 TCGAGAGCGCGACGACGATATCGTGGCGCGGACGATATCGAAGAACAGGAGGTGGCG 186  
 DB 130 TCGATGACGATCCCGACTCTATTACCGTTTGGGAGATTTCGAAGAACAGGAGGTGCC 189  
 QY 187 GTGGGACCTGGAAGCGCAAGGTGCTGTCGACATCCGCGAGTTCTACGAGAGGACGGC 246  
 DB 190 GTGAGGAAGTGGAAAGGAGGAGCATATGTTGACATTCGGGAGTTTACGTCGAAGATGGC 249  
 QY 247 AAGACCCCTCCCGCGCGCAAGGTATACAGCTCCCAATGATCAGTGGAGATACCTGAGG 306  
 DB 250 AAGCAATTGCTGGCAGAAAGGTATCTCTTTGACCATGGATCAGTGGATGTGCTCGT 309  
 QY 307 GACATATCAAGCTATAGATGAGGCCATCAGGAGAAATCGGTGATCGGAGCCCATCT 365  
 DB 310 AATCATGTTGAAGAAATTGCAAGGCAATTAATGAGAAATCTTAGGAAGTGGGCAATCT 368  
 RESULT 9  
 AAC38126  
 ID AAC38126 standard; DNA; 471 BP.  
 XX  
 AC AAC38126;  
 XX  
 DT 17-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 19870.  
 XX  
 KW Hybridisation assay; Genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway; metabolic pathway;  
 KW promoter; termination sequence; ss.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN BP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-00301439.  
 XX  
 XX 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-013049P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137282P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139452P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140931P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142300P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.

PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144633P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149802P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 13-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.

```

PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 21-OCT-1999; 99US-0160880P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 15.2%; Score 96; DB 3; Length 471;
Best Local Similarity 66.3%; Pred. No. 6.8e-15;
Matches 138; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 142 GACGATATGTCGTCCTCCCGAGATATCGAAGAACAGAGGTCGCGTGGACCTGGAC 201
DB 220 GACGATATGTCGTCGTCGCAACATCTAAGATATGAGAGATCTCTGTAAGGAATTGGAAC 279

QY 202 GCGAAGTCTGTCGACATCGCGAGTCTACGAGAACGCGCAAGACCCCTCCCGGC 261
DB 280 GGGAAATTTGGATTGACATTCGTGAGTTCTATGTCAGAGCGGAAGACTTTGCTTGGC 339

QY 262 CGCAAGGTATACAGTCTCCCAATGGATCAGTGGAGATCTAGGGACAAATATCAAAGCT 321
DB 340 AAGAAAGGTATCTCTCTAAGGTGGATCAGTGGACACTCTTCGGAACACCGCAGAGAT 399

QY 322 ATAGATGAGCCATCAAGGAAATGCGT 349
DB 400 ATCGAAGGCCCTCTCTGACCTTTCTT 427

RESULT 10
AAZ50416
ID AAZ50416 standard; cDNA; 516 BP.
XX AC
XX AAZ50416;
XX 18-MAY-2000 (first entry)
XX Vernonia transcription coactivator PC4 (P15) type 1 cDNA.
DE DE
XX Positive cofactor 4; PC4; transcription coactivator; PC4 (P15) type 1;
KW KW
KW Vernonia; signal mediator; activator; Gal4/VP16; transgenic plant;
KW general transcription factor; GTF; transcription initiation complex;
KW immunological screening; detection; marker; clone vsln.pk013.f21; ss.
XX XX
OS Vernonia mespilifolia.
XX Key Location/Qualifiers
FT mat_peptide 1..333
FT /tag= a
FT /product= "Vernonia PC4 (P15) type I transcription
FT coactivator"
FT

```

```

FT XX
PN XX
XX XX
PD XX
XX XX
PF 21-JUL-1999; 99WO-US016479.
PR 22-JUL-1998; 98US-0093687P.
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX Cahoon RE, Caimi PG, Odell JT, Sakai H, Zhu Q;
PI WPI; 2000-182701/16.
XX P-PSDB; AAY44883.
DR
XX
XX Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT used to alter the level of PC4 (P15) type 1 and PC4 (P15) type 2
PT polynucleotides and polypeptides.
XX
XX Claim 3; Page 38; 43pp; English.
XX The present sequence is the cDNA encoding vernonia Positive Cofactor 4
CC (PC4) transcription coactivator, designated as PC4 (P15) type 1. It is
CC isolated from clone vsln.pk013.f21, obtained from vsln cDNA library,
CC prepared from vernonia seeds. PC4 molecules functions as a signal
CC mediator between activators like Gal4/VP16 and general transcription
CC factors (GTFs) in a transcription initiation complex. The expression
CC levels of PC4 can be manipulated and the functional properties of
CC specific transcriptional activators can be modulated. The PC4 DNA can be
CC used to create transgenic plants with altered PC4 levels, that would
CC affect the level of transcription of specific genes in the plant. It is
CC also used for immunological screening of cDNA libraries and to raise
CC specific antibodies for detection. The DNA is used as probes and primers,
CC for genetic and physical mapping of genes and as markers for traits
CC linked to those genes
XX
SQ Sequence 516 BP; 153 A; 103 C; 121 G; 126 T; 0 U; 13 Other;

Query Match 12.4%; Score 78.6; DB 3; Length 516;
Best Local Similarity 60.6%; Pred. No. 2.4e-10;
Matches 120; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 135 CCGCCAGCAGCATATCGTCGTCGCCAGATATCGAAGACAGAGGTCGCGTGGAC 194
DB 129 CGACGAAGACGACATCTTCATTTGTGACGTTTCCAAAACCGGAGGTTTGTTCAGGA 188

QY 195 CTGGAAACGGCAAGTCTGTCGACATCGCGAGTTTCTACGAGAGGACGGCAAGCCCT 254
DB 189 NTGGCAAGGGAGGTCCTTGTGATATCGTNGTGTTCATGAAACGGCAANNAAT 248

QY 255 CCGCGCGCCAAAGGTATACAGTCTCCCAATGATGATGAGTGAAGATCTGAGGCAATAT 314
DB 249 GCTGTGNCAAAANAGGCATCTCATTTGACATGACCATGATCANTGGAAGAACTCCGTCATGT 308

QY 315 CAAAGCTATAGATGAGGC 332
DB 309 GGATGAANTGACAAGGC 326

RESULT 11
AAZ50418
ID AAZ50418 standard; cDNA; 939 BP.
XX AC
XX AAZ50418;
XX 18-MAY-2000 (first entry)
XX Corn transcription coactivator PC4 (P15) type 2 cDNA from contig-2.
DE DE
XX Positive cofactor 4; PC4; transcription coactivator; PC4 (P15) type 2;
KW KW
KW corn; signal mediator; activator; Gal4/VP16; transgenic plant; contig;
KW

```

/note= "Derived from clone vsln.pk013.f21"





```
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 08-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151085P.
PR 27-AUG-1999; 99US-0151086P.
PR 27-AUG-1999; 99US-0151087P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159233P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.

PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 11.7%; Score 74; DB 3; Length 1063;
Best Local Similarity 73.1%; Pred. No. 5e-09;
Matches 95; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 142 GACGATATCGTCTGCGCCGAGATATCGAAGAACAGAGGGTGGCGGTGCGGACTGGAAC 201
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 220 GACGATATCGTCTGCGCAACATATCTAAGATAGGAGAGTCTCTGTAAGGAATTGGAAC 279
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 202 GGCAAGTTCGTCTGACATCCGCGAGTTCACGAGAAGGACCGCAAGACCTCCCGGC 261
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 280 GGGAGATTGGATTGCATTCGTGAGTCTATGTCAGAGCGAAGAGACTTTGCCCTGGC 339
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 262 CGCAAAGGTA 271
DB |||||||
QY 340 AAGAAAGGTA 349

RESULT 14
ACF03526
ID ACF03526 standard; cdna; 913 BP.
XX
AC ACF03526;
XX
DT 27-OCT-2003 (revised)
DT 12-SEP-2003 (first entry)
XX
DE Brassica campestris MIP102 encoding cdna SEQ ID NO:1.
XX
KW Brassica campestris; plant; MIP102; virus resistance;
KW plant virus transport protein binding protein; gene; ss.
XX
OS Brassica rapa.
XX
FH Key Location/Qualifiers
FT CDS 14..511
FT /*tag= a
FT /product= "MIP102"
XX
PN WO2003022039-A1.
XX
PD 20-MAR-2003.
XX
PF 10-SEP-2001; 2001WO-JP007858.
XX
PR 10-SEP-2001; 2001WO-JP007858.
XX
PA (NAAG-) NAT INST AGROBIOLOGICAL SCI.
XX
PI Nishiguchi M, Niyunoya H, Matsushita Y;
XX
WPI; 2003-313169/30.
P-PSDB; ABR57409.
XX
Plant protein binding to plant virus transport protein for imparting
virus resistance to plants including tobacco.
XX
Claim 3; Fig 2; 8app; Japanese.
XX
The present invention describes a method for imparting virus resistance
to plants in which the plant cells are transformed with a polynucleotide
encoding a protein (I) which binds to plant virus transport protein. Also
described are plants transformed by (I). The method is useful for
increasing resistance of plants such as tobacco, tomato, petunia, pear,
rice, wheat, barley, maize, soybean, oilseed rape, rose, apple, alfalfa,
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 11, 2005, 18:41:07 ; Search time 72 Seconds  
(without alignments)  
6239.147 Million cell updates/sec

Title: US-10-629-953-3  
Perfect score: 632  
Sequence: 1 ctctccaagctccacag.....taaaaaaaaaaaaaaaaaa 632

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*

2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*

3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*

4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*

5: /cgn2\_6/ptodata/1/ina/BCTUS.COMB.seq:\*

6: /cgn2\_6/ptodata/1/ina/backfiles.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	182.2	28.8	289	4	US-09-313-294A-1459
2	149.6	23.7	263	4	US-09-313-294A-1612
3	145.6	23.0	286	4	US-09-313-294A-6916
4	55.4	8.8	7218	1	US-08-232-463-14
5	52.4	8.3	1074	3	US-09-248-335-67
6	45.8	7.2	954	4	US-09-252-991A-2562
7	45.8	7.2	1953	4	US-09-252-991A-2481
8	45.8	7.2	2907	4	US-09-252-991A-2354
9	45	7.1	860	3	US-09-248-335-41
10	45	7.1	4403765	3	US-09-103-840A-2
11	45	7.1	4411529	3	US-09-103-840A-1
12	43.4	6.9	1623	4	US-09-252-991A-15700
13	43.4	6.9	2208	4	US-09-252-991A-15669
14	43.2	6.8	2460	4	US-09-252-991A-2998
15	43.2	6.8	2856	4	US-09-252-991A-2869
16	43.2	6.8	3387	4	US-09-252-991A-3101
17	42.8	6.8	675	3	US-09-998-416-723
18	42	6.6	49377	1	US-08-764-233A-1
19	42	6.6	4403765	3	US-09-103-840A-2
20	42	6.6	4411529	3	US-09-103-840A-1
21	41.4	6.6	1395	4	US-09-252-991A-13078
22	41.4	6.6	2157	4	US-09-252-991A-12861
23	41.2	6.5	768	4	US-09-252-991A-1114
24	41.2	6.5	924	4	US-09-252-991A-969
25	41.2	6.5	939	4	US-09-252-991A-925
26	41	6.5	1109	3	US-09-178-610-1
27	40.8	6.5	832	4	US-09-621-976-2813

Sequence 7, Appli  
Sequence 9083, Ap  
Sequence 45, Appl  
Sequence 9253, Ap  
Sequence 8795, Ap  
Sequence 8696, Ap  
Sequence 7, Appli  
Sequence 1, Appli  
Sequence 9282, Ap  
Sequence 9272, Ap  
Sequence 16141, A  
Sequence 64, Appl  
Sequence 18462, A  
Sequence 11, Appl  
Sequence 76, Appl  
Sequence 15011, A  
Sequence 12248, A  
Sequence 12362, A

28 40.8 6.5 2943 1 US-08-042-747A-7  
c 29 40.6 6.4 708 4 US-09-252-991A-9083  
31 40.6 6.4 840 3 US-09-248-335-45  
c 32 40.6 6.4 1221 4 US-09-252-991A-9253  
33 40.6 6.4 1749 4 US-09-252-991A-8795  
34 40.6 6.4 2313 4 US-09-252-991A-8696  
c 35 40.4 6.4 44377 2 US-08-804-227C-7  
36 40.2 6.4 44377 2 US-08-804-198-1  
37 40.2 6.4 3729 4 US-09-252-991A-9272  
38 40 6.3 1134 4 US-09-252-991A-16141  
39 40 6.3 1332 4 US-09-252-991A-16462  
40 40 6.3 2157 4 US-08-624-447-11  
c 41 40 6.3 4052 4 US-09-266-965-76  
42 40 6.3 53500 4 US-08-624-447-11  
43 39.8 6.3 414 4 US-09-252-991A-15011  
44 39.8 6.3 678 4 US-09-252-991A-12248  
c 45 39.8 6.3 969 4 US-09-252-991A-12362

## ALIGNMENTS

### RESULT 1

US-09-313-294A-1459  
; Sequence 1459, Application US/09313294A  
; Patent No. 6476212  
; GENERAL INFORMATION:  
; APPLICANT: Lalngudi, Raghunath V.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Sherman, Bradley K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
; FILE REFERENCE: PL-0017 US  
; CURRENT APPLICATION NUMBER: US/09/313,294A  
; CURRENT FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 7600  
; SOFTWARE: PERL Program  
; SEQ ID NO 1459  
; LENGTH: 289  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. 6476212 700550983H1  
US-09-313-294A-1459

Query Match 28.8%; Score 182.2; DB 4; Length 289;  
Best Local Similarity 79.6%; Pred. No. 9.3e-44;  
Matches 230; Conservative 0; Mismatches 53; Indels 6; Gaps 1;

QY 41 GAAGAATGTGGCGAAGGGGAAACAAGCGGTTCGGCGCGCGCGGC-----GAGCGCGCG 94  
DB 1 GGAAGATGTGGCGAAGGAAAGAGCGTTTCGGCGGTGGCGCGCGGTGAGCGCGCG 60

QY 95 CCAAGCGCGGTTCGGCGCGCGCGCGCGCGCGCTTCGAGAGCGCGCGACGACGATATCGTCG 154  
DB 61 CCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCGAGAGCGCGCGTACCGTCG 120

QY 155 TCGCCCGAGATATCGAAGACAGAGGGTGGCGGTGGCGACCTGGAACGGCGAGGTCTGCG 214  
DB 121 TAGCCGAGATATCGAAGACAGAGGGTGGCGGTGGCGACCTGGAAGCGCGGTCTTCG 180

QY 215 TCACATATCGCGAGTTCTACGAGAGGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 274  
DB 181 TCACATATCGCGAGTTCTACGAGAGGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240

QY 275 AGCTCCCAATGATCGAGTGAAGATATCTGAGGAGCAATATCAAGGCTAT 323  
DB 241 CACTCCCAATGATCGAGTGAAGATATCTGAGGAGCAATATCAAGGCTAT 289

### RESULT 2

US-09-313-294A-1612

```
; Sequence 1612, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalugudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 1612
; LENGTH: 263
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700551220H1
; NAME/KEY: unsure
; LOCATION: 25, 181
; OTHER INFORMATION: a, t, c, g, or other
;
US-09-313-294A-1612

Query Match      23.7%; Score 149.6; DB 4; Length 263;
Best Local Similarity 78.0%; Pred. No. 3.1e-34;
Matches 206; Conservative 0; Mismatches 51; Indels 7; Gaps 2;

QY 41 GAAGATGTGGCGAAGGGAACACGCGGTTCGGCGCGCGCGCG 94
DB 1 GGAAGATGTGGGGAAGGAAGGCGTTTCGGCGGTGGCGCGCGCG 60
QY 95 CCAAGCGCGTGGCGCGGAGCAGCGGCCCTCCGAGAGCGCGCGATATCGTCG 154
DB 61 CCAAGCGCGGAGCGCGGAGGAGATGCGCCCTCCGATCCGCTGAAGCGGTACCGTCG 120
QY 155 TCGCCCGAGATATCGAAGAACAGGAGGTTGGCGGTGGCGACCTCGAAGCGCAAGGTGTCG 214
DB 121 TAGCCGAGATATCGAAGAACAGGAGTGTCCGTTAGGAGCTGGAAAGCGAGGTCTTCG 180
QY 215 TCGACATCCGCGAGTTCTAGAGAGGAGCAGCGAAGACCTCCCGCGCGCAAAAGGTATAC 274
DB 181 NCGACTTACGCGAGTTCTACTTCAAGGACGCGAAGACTCTCCCGACCGCAAAAGGTATA - 239
QY 275 AGCTCCCAATGGATCAGTGAAGA 298
DB 240 TCATCCANTAGATCAGTGAAGA 263

RESULT 3
US-09-313-294A-6916
; Sequence 6916, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalugudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 6916
; LENGTH: 286
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700380835H1
; NAME/KEY: unsure
; LOCATION: 54
; OTHER INFORMATION: a, t, c, g, or other
;

US-09-313-294A-6916

Query Match      23.0%; Score 145.6; DB 4; Length 286;
Best Local Similarity 78.4%; Pred. No. 4.9e-33;
Matches 189; Conservative 0; Mismatches 45; Indels 7; Gaps 1;

QY 40 AGAAGATGTGGCGAAGGGAACAAAGCGGTTTC-----GGCGCGCGCGCGCGCGCG 92
DB 44 AGGAAGATGTGGGGAAGGAAGAAAGCGTTTCGGCGGTGGCGCGCGCGCGCGCGCG 103
QY 93 GGCCTAAGCGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 152
DB 104 GGCCTAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 163
QY 153 CGTCGCGCGCGAGATATCGAAGAACAGGAGGTTGGCGGTGGCGCGCGCGCGCGCGCG 212
DB 164 CGTAGCGGAGATATCGAAGAACAGGAGGTTTCGTTAGGAGCTGGAAAGCGAGGTCTT 223
QY 213 CGTCGACATCCGCGAGTTCTTACGAGAACGAGCGCAAGACCTCCCGCGCGCGCAAGGTAT 272
DB 224 CGTCGACTTACGCGAGTTCTTACTTCAAGGACGCGCAAGACTCTCCCGACCGCAAGGTAT 283
QY 273 A 273
DB 284 A 284

RESULT 4
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 883-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
```

```
; CLONE: PTZgpt-F1s
US-08-232-463-14

Query Match      8.8%; Score 55.4; DB 1; Length 7218;
Best Local Similarity 6.7%; Pred. No. 7.5e-06;
Matches 29; Conservative 223; Mismatches 179; Indels 0; Gaps 0;

Qy 11 ACTCCACAAGTCCAGAGGAGAAAGTGTGAGAGAAATGTGCGGAGGGAACAAAGCGGT 70
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1436 ACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1377

Qy 71 TCGCGCGCGCGGAGCGCGCGCAAGCCCTGCGCGCGGAGCAGCGGCCCTCCG 130
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1376 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1317

Qy 131 AGAGCGCGGACGACGATATCTGTCGCGCCACGATATCAAGAACAGGAGGTGCGGTGC 190
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1316 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1257

Qy 191 GGACTGGAACGCGAAGTCTGTCGACATCCCGAGTTCACGAGAAGCAGCGCAAGA 250
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1256 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1197

Qy 251 CCCTCCCGCGCAAGGTATACAGCTCCCAATGGATCAGTGAAGATCTGAGGACA 310
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1196 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1137

Qy 311 ATATCAAGCTATAGATGAGCCCATCAAGGAGAAATGCGTGATCGAGCCCATCTCTGT 370
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1136 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1077

Qy 371 GATCAAGTAGACAAAGCTACGCTCTGTTTATGATCCGAGGAATATTGCACTTTTG 430
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1076 RRRRRRRRRATCGAAGCTCCCTCGACCTGCAGCCAAGCTCGGAATTAATCTGTGAGC 1017

Qy 431 GTATGCTAATC 441
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1016 GTATGGCAAC 1006

RESULT 5
US-09-248-335-67
; Sequence 67, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEF, DANIEL
; TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
; FILE REFERENCE: CH-1128-A
; CURRENT APPLICATION NUMBER: US/09/248,335
; CURRENT FILING DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: 08/924,759
; EARLIER FILING DATE: 1997-September-05
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 67
; LENGTH: 1074
; TYPE: DNA
; ORGANISM: maize
US-09-248-335-67

Query Match      8.3%; Score 52.4; DB 3; Length 1074;
Best Local Similarity 5.3%; Pred. No. 2e-05;
Matches 116; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

Qy 71 TCGCGCGCGCGGAGCGCGCGCAAGCCCTGCGCGCGGAGCAGCGGCCCTCCG 130
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
91 TCAGGGCGCGGTGCGCTAAACCTCAAGGCCCTGCGCTACCGATACGTGAGGACAACC 150

Qy 131 AGAGCGCGGACGACGATATCTGTCGCCACGATATCGAAGAACAGAGGCTGCGGTGC 190
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
151 TGGACAGCAAGAGCGAGTCTCTCTCGCTCCACCCCGTGACGGGAAGTGCGGTGC 210
```

```
Qy 191 GGACCTGGAAACGGCAAGGTGCTGTCACATCCGCGAGTTCTACGAGAAGACGGCAAGA 250
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
211 TCCTCCACGACGCGGAGCGCGTCTGCGAGTCCCGGTCATCGTGAGTATATCGACGAGG 270

Qy 251 CCCTCCCGCGCGCAAGGTATACAGCTCCCAATGGATCAGT 292
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
271 CTTCCCGCGCAGCGGCCCTGCTCTCTCCCGCGCACCGCT 312

RESULT 6
US-09-252-991A-2562
; Sequence 2562, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2562
; LENGTH: 954
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2562

Query Match      7.2%; Score 45.8; DB 4; Length 954;
Best Local Similarity 53.7%; Pred. No. 0.0016;
Matches 95; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 71 TCGCGCGCGCGCGGAGCGCGCCAGCCGTCGCGCGCGGAGCAGCGGCCCTCCG 130
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
299 TCGCGGACCTGCGCGGTGCGCGCCAGCGCGGTGCGCTACGAGAACCGCGCACCGTGG 358

Qy 131 AGAGCGCGGACGACGATATCTGTCGCCACGATATCGAAGAACAGAGGCTGCGGTGC 190
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
359 AGTTCCTGCTGCGCGGAGCGGTGACTTCATGNGATGACACCCGGGTCCAGGTGG 418

Qy 191 GGACCTGGAAACGGCAAGGTGCTGTCGACATCCCGAGTTCTACGAGAAGACGGCA 247
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
419 AGCACACCATCACCGAGGAATCACCGCATCAGCGTGTCCGCGACGATCCGCA 475

RESULT 7
US-09-252-991A-2481
; Sequence 2481, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2481
; LENGTH: 1953
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2481

Query Match      7.2%; Score 45.8; DB 4; Length 1953;
Best Local Similarity 53.7%; Pred. No. 0.0024;
Matches 95; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
```



## ; GENERAL INFORMATION:

; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; TITLE OF INVENTION: TUBERCULOSIS  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4411529  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: H37Rv  
US-09-103-840A-1

Query Match 7.1%; Score 45; DB 3; Length 4411529;  
Best Local Similarity 50.7%; Pred. No. 0.28;  
Matches 108; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

Qy 58 GGGAAACAAAGCGTTTCGGCGGCGCGAGCCGCGCCAGAGCGCCGTCGCGCGGGAC 117  
Db 925548 GGGCGCGGGGTTCCGCGGTTCCGCGGCGAGCTGTTGGGCGCGTGGCGCGGGGCC 925607  
Qy 118 GACGGCGCCTCCGAGAGCGCGAGCGAGATATCGTCCGCCAGATATCGAAGAACAGG 177  
Db 925608 GCGCGGCTGTTCCGACAGCGCGGCGAGCGGGTTCGCGGTTCCGATCGACGGCACC 925667  
Qy 178 AGGTGGCGGTGCGGACTGGAACGGAAGTCTGTCGATCCGCGGAGTTCTACGAG 237  
Db 925668 GGTGGGCGCGCGCACCGGTGGTGGTGTGTCGACGCGCGGGTTCGGCGGT 925727  
Qy 238 AAGGACGCGAAGACCTCCCGCGCGCAAGGT 270  
Db 925728 ACTGCGGGGAGCGCGGTTCGCGGTACCGGT 925760

## RESULT 12

US-09-252-991A-15700  
; Sequence 15700, Application US/09252991A  
; Patent No. 6551795

## ; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 15700  
; LENGTH: 1623  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-15700

Query Match 6.9%; Score 43.4; DB 4; Length 1623;  
Best Local Similarity 51.9%; Pred. No. 0.011;  
Matches 98; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

Qy 78 CGGCGGAGCGCGCGCCAGCCCTGCGCGCGGAGCAGCGGCCCTCCGAGCGC 137  
Db 1005 CGGCGCAAGCCAGCGCTCTCCGCGGTTCATCGCGGAGCGGAGTTGGCCGAGCCGC 1064  
Qy 138 CGACGAGATATCGTCTGCCCGCATATCGAAGAACAGGAGGTGCGGTGCGGACCTG 197  
Db 1065 CGATACCGTGTCCCCACCGCGGAGATCTTCCGCTACTGGTGCAGGCGGACGATCGA 1124

Qy 198 GAACGGCAAGGTCTGTCGATCCCGAGTTCTACGAGAAGACGGCAAGACCCCTCCC 257  
Db 1125 CGTCGGCTTCTCGCGCGCGCCAGGTTCGACCGCTTCGGCAACATCAACACCGGTGAT 1184  
Qy 258 CGGCGCGCAA 266  
Db 1185 CGGCGACTA 1193

## RESULT 13

US-09-252-991A-15669  
; Sequence 15669, Application US/09252991A  
; Patent No. 6551795

## ; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 15669  
; LENGTH: 2208  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-15669

Query Match 6.9%; Score 43.4; DB 4; Length 2208;  
Best Local Similarity 51.9%; Pred. No. 0.013;  
Matches 98; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

Qy 78 CGGCGGAGCGCGCGCCAGCGCTGCGCGCGGAGCAGCGGCCCTCCGAGCGGC 137  
Db 388 CGGCGCAAGCCAGCGTCTCTCGCGTCTCATCGCGGAGTTGGCCGAGCCGC 447  
Qy 138 CGACGAGATATCGTCTGCCCGCATATCGAAGAACAGGAGGTGCGGTGCGGACCTG 197  
Db 448 CGATACCGTGTCCCGACCGCGGAGATCTTCCGCTACTGCGGCGGACGATCGA 507  
Qy 198 GAACGGCAAGGTCTGTCGATCCCGAGTTCTACGAGAAGACGGCAAGACCCCTCCC 257  
Db 508 CGTCGGCTTCTCGCGCGCGCCAGGTTCGACCGCTTCGGCAACATCAACACCGGTGAT 567  
Qy 258 CGGCGCGCAA 266  
Db 568 CGGCGACTA 576

## RESULT 14

US-09-252-991A-2998  
; Sequence 2998, Application US/09252991A  
; Patent No. 6551795

## ; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 2998  
; LENGTH: 2460  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa

## US-09-252-991A-2998

Query Match 6.8%; Score 43.2; DB 4; Length 2460;  
Best Local Similarity 50.5%; Pred. No. 0.016;  
Matches 105; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 86 AGCGCGCGGCAAGCGCGGTGCGCGCGGAGCAGCGGCCCTCCGAGAGCGCGCGACG 145  
DB 1902 ACCCGTGGCGGAGCGGATGACCGGTGGACCTTGCAGGAGCCATCGGCGCGCCACG 1961

QY 146 ATATCGTGTGCGCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAAACGGCA 205  
DB 1962 AACTGGTCTTCAACGTCTCGACAGCGACACCGCGGAGGCCCTGGAAAGCCCGTGCAC 2021

QY 206 AGTGTGTGTGCGATCCGCGAGTTCTACGAGAGGACGGCAGACCCTCCCGCGCGCA 265  
DB 2022 GCTGCCCTGCAGAACACCGCGGTGTCTACTGGAGGACGCGCCACGCTGATCAGCGCA 2081

QY 266 AAGGTATACAGTCCCAATGGATCAGTG 293  
DB 2082 AGGCGAGCAGCAGGATATCCAGNATTG 2109

## RESULT 15

US-09-252-991A-2869/c  
; Sequence 2869, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 2869  
; LENGTH: 2856  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-2869

Query Match 6.8%; Score 43.2; DB 4; Length 2856;  
Best Local Similarity 50.5%; Pred. No. 0.017;  
Matches 105; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 86 AGCGCGCGGCAAGCGCGGTGCGCGCGGAGCAGCGGCCCTCCGAGAGCGCGCGACG 145  
DB 1625 ACCCGTGGCGGAGCGGATGACCGGTGGACCTTGCAGGAGCCATCGGCGCGCCACG 1566

QY 146 ATATCGTGTGCGCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAAACGGCA 205  
DB 1565 AACTGGTCTTCAACGTCTCGACAGCGACACCGCGGAGGCCCTGGAAAGCCCGTGCAC 1506

QY 206 AGTGTGTGTGCGATCCGCGAGTTCTACGAGAGGACGGCAGACCCTCCCGCGCGCA 265  
DB 1505 GCTGCCCTGCAGAACACCGCGGTGTCTACTGGAGGACGCGCCACGCTGATCAGCGCA 1446

QY 266 AAGGTATACAGTCCCAATGGATCAGTG 293  
DB 1445 AGGCGAGCAGCAGGATATCCAGNATTG 1418

Search completed: January 11, 2005, 21:22:20  
Job time : 96 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 11, 2005, 21:06:03 ; Search time 483 Seconds  
(without alignments)  
7507.856 Million cell updates/sec

Title: US-10-629-953-3

Perfect score: 632

Sequence: 1 ctctccaaagactccacag.....taaaaaaaaaaaaaaaaaaaaaa 632

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4293498 seqs, 2868903791 residues

Total number of hits satisfying chosen parameters: 8586996

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
- 20: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 21: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	610.8	96.6	1754	US-10-437-963-83372	Sequence 83372, A
2	208	32.9	660	US-10-425-114-4500	Sequence 4500, Ap
3	204.8	32.4	757	US-10-425-115-74342	Sequence 74342, A
4	146.6	23.2	310	US-09-294-0938-67	Sequence 67, Appl
5	109.4	17.3	663	US-10-424-599-38767	Sequence 38767, A
6	102.6	16.2	913	US-10-424-599-38768	Sequence 38768, A
7	76	12.0	933	US-10-425-114-24336	Sequence 24336, A
8	76	12.0	935	US-10-425-114-33158	Sequence 33158, A
9	76	12.0	1225	US-10-425-115-154421	Sequence 154421, A
10	74.8	11.8	658	US-10-425-114-21115	Sequence 21115, A
11	74.8	11.8	658	US-10-425-115-20439	Sequence 20439, A
12	73.2	11.6	471	US-10-767-701-29085	Sequence 29085, A

C 13	67.4	10.7	497	16	US-10-424-599-127795	Sequence 127795, A
14	66.2	10.5	794	16	US-10-425-114-20666	Sequence 20666, A
C 15	66.2	10.5	911	16	US-10-424-599-4356	Sequence 4356, Ap
C 16	64.4	10.2	1030	17	US-10-437-963-59601	Sequence 59601, A
17	63.2	10.0	409	18	US-10-425-115-157838	Sequence 157838, A
18	62.6	9.9	1479	16	US-10-424-599-9157	Sequence 9157, Ap
C 19	61	9.7	496	10	US-09-770-961-520	Sequence 520, App
20	52	8.2	892	18	US-10-425-115-83949	Sequence 83949, A
C 21	50.8	8.0	1168	18	US-10-425-115-150589	Sequence 150589, A
C 22	50.6	8.0	1887	17	US-10-437-963-75128	Sequence 75128, A
23	49.4	7.8	1651	17	US-10-767-701-13956	Sequence 13956, A
C 24	49.2	7.8	900	16	US-10-425-114-22315	Sequence 22315, A
25	49.2	7.8	1015	16	US-10-425-114-25767	Sequence 25767, A
C 26	48.8	7.7	647	16	US-10-425-114-16885	Sequence 16885, A
27	48.8	7.7	2209	17	US-10-437-963-69229	Sequence 69229, A
C 28	48	7.6	1996	17	US-10-437-963-46655	Sequence 46655, A
29	47.8	7.6	542	17	US-10-767-701-3376	Sequence 3376, Ap
C 30	47.8	7.6	693	17	US-10-437-963-68177	Sequence 68177, A
31	47.6	7.5	597	16	US-10-260-238-5561	Sequence 5561, Ap
C 32	47.6	7.5	1071	17	US-10-437-963-94451	Sequence 94451, A
33	47.6	7.5	1397	16	US-10-425-114-19393	Sequence 19393, A
C 34	47.6	7.5	1755	16	US-10-425-114-904	Sequence 904, App
35	47.6	7.5	1755	18	US-10-425-115-100553	Sequence 100553, A
C 36	47.6	7.5	1810	16	US-10-425-114-31485	Sequence 31485, A
37	47.6	7.5	2365	18	US-10-425-115-100551	Sequence 100551, A
C 38	47.4	7.5	2256646	17	US-10-470-565-1	Sequence 1, Appli
39	47.2	7.5	1509	10	US-09-967-464-68	Sequence 68, Appl
C 40	46.8	7.4	1848	16	US-10-293-913A-1	Sequence 1, Appli
41	46.8	7.4	1848	16	US-10-293-913A-3	Sequence 3, Appli
C 42	46.8	7.4	3123	18	US-10-425-115-85044	Sequence 85044, A
43	46.6	7.4	1497	16	US-10-425-114-24433	Sequence 24433, A
C 44	46.6	7.4	1535	18	US-10-425-115-67209	Sequence 67209, A
45	46.4	7.3	1509	10	US-09-967-464-64	Sequence 64, Appl

#### ALIGNMENTS

#### RESULT 1

US-10-437-963-83372  
; Sequence 83372, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 83372  
; LENGTH: 1754  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_8270C.1  
US-10-437-963-83372

Query Match 96.6%; Score 610.8; DB 17; Length 1754;  
Best Local Similarity 99.7%; Pred. No. 1.7e-179;  
Matches 612; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CTCTCCAAAGCTCCACAGTCCAGGGGAGAGTGTGAGAGAAATGTGCGGAGGGG 60  
Db 1078 CTCTCCAAAGCTCCACAGTCCAGGGGAGAGTGTGAGAGAAATGTGCGGAGGGG 1137

QY 61 AACAAAGCGTTTCGGCGCGCGCGAGCCGCGCGCAAGCGCGCTGCGCGCGGAGCGAC 120  
Db 1138 AACAAAGCGTTTCGGCGCGCGCGAGCCGCGCGCAAGCGCGCTGCGCGCGGAGCGAC 1197  
QY 121 GGGCCCTCCGAGAGCGCGGAGAGAGATATCGTGTGCGCCGAGATATCGAAGAAACAGAGG 180  
Db 1198 GGGCCCTCCGAGAGCGCGGAGAGAGATATCGTGTGCGCCGAGATATCGAAGAAACAGAGG 1257  
QY 181 GTGCGGTGCGGACCTCGAAGCGCAAGGTGCGTGCGACATCCGCGAGTTCACGAGAAG 240  
Db 1258 GTGCGGTGCGGACCTCGAAGCGCAAGGTGCGTGCGACATCCGCGAGTTCACGAGAAG 1317  
QY 241 GACGGCAAGACCTCCCGCGCGCAAGGTATACAGCTCCCAATGGATCACTAGTGGAGATA 300  
Db 1318 GACGGCAAGACCTCCCGCGCGCAAGGTATACAGCTCCCGATGGATCACTAGTGGAGATA 1377  
QY 301 CTGAGGAGCAATCAAGCTATAGATGAGGCGCATCAAGAGAAATGCGTGATCGGAGCCC 360  
Db 1378 CTGAGGAGCAATCAAGCTATAGATGAGGCGCATCAAGAGAAATGCGTGATCGGAGCCC 1437  
QY 361 ATTCTCTGTGATGCAAGTAGACTAAGCCTACGCTCTGCTTTTATGACTCCGAGGAATAT 420  
Db 1438 ATTCTCTGTGATGCAAGTAGACTAAGCCTACGCTCTGCTTTTATGACTCCGAGGAATAT 1497  
QY 421 TGCACTTTTGGTATGTAATCTTTTATTAACCTGCTAGTAAATATCAAGTAATGTTTTAGT 480  
Db 1498 TGCACTTTTGGTATGTAATCTTTTATTAACCTGCTAGTAAATATCAAGTAATGTTTTAGT 1557  
QY 481 TTGGTCAATGCGGAAATGTGCGCTTGTCTGTATTTGTCGCAAGAGACGAGCTATGTGC 540  
Db 1558 TTGGTCAATGCGGAAATGTGCGCTTGTCTGTATTTGTCGCAAGAGACGAGCTATGTGC 1617  
QY 541 CTAATGTGCTCTCTCAATTTGTACGTTGGTAAGGAGGTCATCTGTTAATGTTGGTT 600  
Db 1618 CTAATGTGCTCTCTCAATTTGTACGTTGGTAAGGAGGTCATCTGTTAATGTTGGTT 1677  
QY 601 GGTTCCTGCGTGA 614  
Db 1678 GGTTCCTGCGTGA 1691

## RESULT 2

US-10-425-114-4500  
; Sequence 4500, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaka, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 4500  
; LENGTH: 660  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700380835\_FLI  
US-10-425-114-4500

Query Match 32.9%; Score 208; DB 16; Length 660;  
Best Local Similarity 79.8%; Pred. No. 5e-54;  
Matches 260; Conservative 0; Mismatches 60; Indels 6; Gaps 1;  
QY 40 AGAAGATGTGGCGGAGGAGCAAGCGGTTGCGGCGCGCGC-----GAGCGCGCG 93  
Db 45 AGGAAGATGTGGCGGAGGAGCAAGCGGTTGCGGCGCGCGCGCGCGCGCGCGCGCG 104

QY 94 GCCAAGCGCGTTCGCGCGGAGCGACGCGCCCTCCGAGAGCGCGACGACGATATCGTC 153  
Db 105 GCCAAGCGCGCGCGGAGGACGATGCGCCCTCCGATCCGATCCGCTGAGACGCTACCGTC 164  
QY 154 GTCGCCAGATATCGAAGAACAGAGGGTGGCGGTGCGGACCTGGAAACGCGCAAGGTGTC 213  
Db 165 GTAGCGAGATATCGAAGAACAGAGGTGTCCGTTAGGAGCTGGAAAGCGCGGCTCTTC 224  
QY 214 GTCGACATCGCGAGTTCACGAGAGGAGCGCAAGCCCTCCCGCGCGCAAGGTATA 273  
Db 225 GTCGACTTACGCGAGTTCCTTCAAGSAGCGCAAGACTCTCCCAACCGCAAGGTATA 284  
QY 274 CAGCTCCCAATGGATCAGTGGAAAGTACTGAGGCAATATCAAAGCTATAGATGAGGCC 333  
Db 285 TCACTCCATTAGATCAGTGGAGATATTGAGGACAAACATCAAGCCATTAATGAGGCA 344  
QY 334 ATCAAGAGAAATGCGTATCGGAGCC 359  
Db 345 ATCGAGGAAATACATGATGGGAGCC 370

## RESULT 3

US-10-425-115-74342  
; Sequence 74342, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 74342  
; LENGTH: 757  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_167799C.1  
US-10-425-115-74342

Query Match 32.4%; Score 204.8; DB 18; Length 757;  
Best Local Similarity 74.8%; Pred. No. 5.3e-53;  
Matches 256; Conservative 0; Mismatches 87; Indels 0; Gaps 0;  
QY 17 CAAGTCCAGAGGAGAAAGTGTGAGAAAGTGTGGCGGAGGAGGAAACAAAGCGGTTCGGCG 76  
Db 49 CCAGTCCAGTCTCAGCAAGAGAGATGTGGGAGAGAGAAAGAGCGTTTCGGCGGTGGCG 108  
QY 77 GCGCGCGGAGCGCGCGCAAGCGCGCTGCGCGCGGAGACGACGCGGCTCCGAGAGCG 136  
Db 109 GCAGCGGTGAGCGCGCGCGCAAGCGCGCAGCGGAGGAGCGATGCGCCCTCCGAATCCG 168  
QY 137 CCGACGACGATATCGTGTGCGCCAGATATCGAAGAACAGAGGTTGGCGGTGCGGACCT 196  
Db 169 CTGAAGACGGTACCGCTCGTAGCGAGATATCGAAGAACAAAGAGGTGTTCGTTAGGAGCT 228  
QY 197 GGAACGGCAAGGTGCTGCTCGACATCCGCGAGTTCACGAGAGGACGCGCAAGACCTCC 256  
Db 229 GGAAGCGAGGCTCTTCGTCGACTTACGCGAGTTCCTACTTCAAGGAGCGCAAGACTCTCC 288  
QY 257 CCGCGCGCAAGGTATACAGCTCCCAATGGATCAGTGGAGATATCTGAGGAGCAATATCA 316  
Db 289 CCACCGCGCAAGGTATATCACTCCCAATTAGATCAGTGGAGATATTGAAAGCAACATCA 348  
QY 317 AGCTATAGATGAGGCGCATCAAGGAAATGCGTGATCGGAGCC 359  
Db 349 AAGCCATTAATGAGGCCATCGAGGAAATACATGATGGTTGCC 391



## RESULT 4

US-09-294-093B-67  
; Sequence 67, Application US/09294093B  
; Patent No. US20010051335A1  
; GENERAL INFORMATION:  
; APPLICANT: Ito, Laura, Y.  
; APPLICANT: Sherman, Bradley, K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL  
; FILE REFERENCE: PL-0009 US  
; CURRENT APPLICATION NUMBER: US/09/294,093B  
; CURRENT FILING DATE: 1999-04-16  
; PRIOR APPLICATION NUMBER: 60/082,567  
; PRIOR FILING DATE: April 21, 1998  
; NUMBER OF SEQ ID NOS: 6207  
; SOFTWARE: PERL Program  
; SEQ ID NO 67  
; LENGTH: 310  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700281907H2  
; NAME/KEY: unsure  
; LOCATION: 2, 30, 295  
; OTHER INFORMATION: a, t, c, g, or other  
US-09-294-093B-67

Query Match 23.2%; Score 146.6; DB 9; Length 310;  
Best Local Similarity 71.2%; Pred. No. 4.9e-35;  
Matches 220; Conservative 0; Mismatches 86; Indels 3; Gaps 2;

Qy	17	CAAGTCAGAGGAGAAAGTGTGAGAAATGTGGCGAAGGGAACAAGCGGTTCGGCG	76
Db	3	CCAGTCAGTCTCAGCAGAGNAGATGTGGGAGGGAAGAGCGTTTCGGCGGTGGCG	62
Qy	77	GGCGCGCGAGCGCGCGCAAGCGCGTGGCGCC-GGGAGCAGCGGCGCTCCGAGAGC	135
Db	63	GCAGCGCGAGCGCGCGCAAGCGCGGAGGAGCGGAGCGATGGCGCCCTCCGAATCC	122
Qy	136	GCCGACGACGATCGTGTGCGCCAGATATCGAAGACAGAGGTTGGCGTGGCGACC	195
Db	123	GCTGAAGCGGTACCGTGTGCGCGAGATATCGAAGAACAGAGGTATCCGTTAGAGC	182
Qy	196	TGGAACGGCAAGTGTGTCGACATCGCGAGTTCACGAGAGGACGGCAAGACCTC	255
Db	183	TGGAAGCGAGGTCTTGTGCTTACGAGTTCCTTCAAGGAGCGGAGACTCTC	242
Qy	256	CCCGCGCGCAAGGTATACAGTCCCAATGGATCAGTGGAGATACCTGAGGACAATATC	315
Db	243	CCCAACCGCAAGGTATATCA--TTCCAATTAGATCATGGAAGATATTGAAGGACNACATC	300
Qy	316	AAAGCTATA 324	
Db	301	AAAGCCATA 309	

## RESULT 5

US-10-424-599-38767  
; Sequence 38767, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 38767  
; LENGTH: 663  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_135006C.1  
US-10-424-599-38767

Query Match 17.3%; Score 109.4; DB 16; Length 663;  
Best Local Similarity 66.1%; Pred. No. 2.9e-23;  
Matches 159; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

Qy	127	TCGAGAGCGCCGACGACGATATCGTGTGCGCCAGATATCGAAGACAGAGGGTGGCG	186
Db	259	TCGATGACGATCCCGACTCTATTACCGTTTGGGAGATTTGGAAGACAGAGGGTTGCC	318
Qy	187	GTGCGACCTGGAAAGCGGAGGTGCTGTCGACATCGCGAGTTCTACGAGAAGAGCGGC	246
Db	319	GTGAGGAACTGGAAGGCGACATTTATGTTGACATTCGCGAGTTTACGTCAAGATGGC	378
Qy	247	AAGACCTTCCCGCGCAAGGTATACAGCTCCCAATGGATCAGTCGAAGATACTGAGG	306
Db	379	AAGCAATTGCTGGCAAGAAAGGTATCTCTTTGACCATGATCAGTGGAAATGTCTCGT	438
Qy	307	GACAAATCAAGCTATAGATGAGGCCATCAAGAGAAATCGGTGATCGGAGCCCATTTCT	365
Db	439	AATCATGTTGAAGAAATTGCAAGGCAATTAATGAGAAATCTTAGGAAGTGGGCAATCT	497

## RESULT 6

US-10-424-599-38768  
; Sequence 38768, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 38768  
; LENGTH: 917  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_135007C.1  
US-10-424-599-38768

Query Match 16.2%; Score 102.6; DB 16; Length 917;  
Best Local Similarity 62.1%; Pred. No. 4.5e-21;  
Matches 162; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

Qy	127	TCGAGAGCGCCGACGACGATATCGTGTGCGCCAGATATCGAAGACAGAGGGTGGCG	186
Db	341	TCGAGACGATCCCGACTCTGTACCATTTTGGAAATTTGGAAGACAGAGGGTTGCC	400
Qy	187	GTGCGACCTGGAAAGCGGAGGTGCTGTCGACATCCGCGAGTTCTACGAGAAGAGCGC	246
Db	401	GTGAGGAACCTGGAAGGCGACATTTATGTTGACATTCGCGAGTTTACGTCAAGATGGC	460
Qy	247	AAGACCTTCCCGCGCAAGGTATACAGCTCCCAATGGATCAGTCGAAGATACTGAGG	306
Db	461	ARGCAATTCCTGGCAGGAAGGTATCTCGTTGACCATGATCAGTGGAAATGTCTCGT	520
Qy	307	GACAAATCAAGCTATAGATGAGGCCATCAAGGAGAAATCGGTGATCGGAGCCCATTTCT	366
Db	521	AATCATGTTGAAGAAATTGACAGGCGAGTTAATGAGAAATCTTAGGAAGTGGGCACTAG	580
Qy	367	TTGTGATGCAAGTAGACTAAG 387	

Db 581 TCAATTTAATAGTTTACTATG 601

RESULT 7

US-10-425-114-24336  
 ; Sequence 24336, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 24336  
 ; LENGTH: 933  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: LIB3608-031-G8\_FLI  
 US-10-425-114-24336

Query Match 12.0%; Score 76; DB 16; Length 933;  
 Best Local Similarity 51.8%; Pred. No. 9.1e-13;  
 Matches 172; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

Qy	25	GAGGAGAAAGTGTGAGAAAGTGTGCGGAGGGAACAAAGCGGTTTCGGCGCGCGGC 84
Db	264	GAGGCGTCTACTGCGAGGAGAGCAAGGGCAAGAACACAGAGGAGGACGAAGAGGAT 323
Qy	85	GAGCGCGCGGCAAGCGCGTGCCTCGCGGAGCAGCGGCGCTCCGAGAGCGCGACGAC 144
Db	324	GACGATGAGGAGGAGGATGAAGTTAAGGTGGCGGAGAGAGTACGATGACCAAGGT 383
Qy	145	GATATCGTCTGCGCCAGATATCGAAGACAGAGGAGGTCGGTGGCGGACCTTGAACGGC 204
Db	384	GACCTTATCTGTGCGCGCTTTCGAGCAAGAGGAGGAGTCTTTATCGGAGTTTAAGGGC 443
Qy	205	AAGTCTCTGTCGACATCCGCGGTTCTACGAGAGGAGCGGACCGCTCCCGCGCGC 264
Db	444	AGTCTGTTGTTCCATCCGCGAGTTCTACGTGAAGAGCGCAAGGAGATGCCCTCCGCC 503
Qy	265	AAAGGTATACAGTCCCAATGGATCAGTGAAGATCTGAGGAGCAATATCAAGCTTATA 324
Db	504	AAAGGTATACAGTATGACTTTGGAGCATGGGAGCAATTTGCAATGCTGTACCTGCAATA 563
Qy	325	GATGAGGCCATCAAGGAGATGCGTGATCGGA 356
Db	564	GAGGATGCCATCAAAAAGCTTGAAGATTTCAGA 595

RESULT 8

US-10-425-114-33158  
 ; Sequence 33158, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 33158  
 ; LENGTH: 935  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: UC-ZMFLMOL17069H06\_FLI  
 US-10-425-114-33158

Query Match 12.0%; Score 76; DB 16; Length 935;  
 Best Local Similarity 51.8%; Pred. No. 9.1e-13;  
 Matches 172; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

Qy	25	GAGGAGAAAGTGTGAGAAAGTGTGCGGAGGGAACAAAGCGGTTTCGGCGCGCGGC 84
Db	282	GAGGCGTCTACTGCGAGGAGAGCAAGGGCAAGAACACAGAGGAGGACGAAGAGGAT 341
Qy	85	GAGCGCGCGGCAAGCGCGTGCCTCGCGGAGCAGCGGCGCTCCGAGAGCGCGACGAC 144
Db	342	GACGATGAGGAGGAGGATGAAGTTAAGGTGGCGGAGAGAGTACGATGACCAAGGT 401
Qy	145	GATATCGTCTGCGCCAGATATCGAAGACAGAGGAGGTCGGTGGCGGACCTTGAACGGC 204
Db	402	GACCTTATCTGTGCGCGCTTTCGAGCAAGAGGAGGAGTCTTTATCGGAGTTTAAGGGC 461
Qy	205	AAGTCTCTGTCGACATCCGCGGTTCTACGAGAGGAGCGGACCGCTCCCGCGCGC 264
Db	462	AGTCTGTTGTTCCATCCGCGAGTTCTACGTGAAGAGCGCAAGGAGATGCCCTCCGCC 521
Qy	265	AAAGGTATACAGTCCCAATGGATCAGTGAAGATCTGAGGAGCAATATCAAGCTTATA 324
Db	522	AAAGGTATACAGTATGACTTTGGAGCATGGGAGCAATTTGCAATGCTGTACCTGCAATA 581
Qy	325	GATGAGGCCATCAAGGAGATGCGTGATCGGA 356
Db	582	GAGGATGCCATCAAAAAGCTTGAAGATTTCAGA 613

RESULT 9

US-10-425-115-154421  
 ; Sequence 154421, Application US/10425115  
 ; Publication No. US20040214272A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(53222)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,115  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 369326  
 ; SEQ ID NO 154421  
 ; LENGTH: 1225  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: MRT4577\_72412C.1  
 US-10-425-115-154421

Query Match 12.0%; Score 76; DB 18; Length 1225;  
 Best Local Similarity 51.8%; Pred. No. 1e-12;  
 Matches 172; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

Qy	25	GAGGAGAAAGTGTGAGAAAGTGTGCGGAGGGAACAAAGCGGTTTCGGCGCGCGGC 84
Db	343	GAGGCGTCTACTGCGAGGAGAGCAAGGGCAAGAACACAGAGGAGGACGAAGAGGAT 402
Qy	85	GAGCGCGCGGCAAGCGCGTGCCTCGCGGAGCAGCGGCGCTCCGAGAGCGCGACGAC 144
Db	403	GACGATGAGGAGGAGGATGAAGTTAAGGTGGCGGAGAGAGATGACGATGACCAAGGT 462

QY 145 GATATCGTCGCGCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGC 204  
|||  
Db 463 GACCTTATCTCTGTCGCGCTTCGAGCAAGAGGAGGTGACTTTATCGGAGTTTAAAGGC 522  
|||  
QY 205 AAGTGTGTCGTCGACATCCGCGAGTCTTACGAGAAGGACGCGCAAGACCTCCCGGGCGC 264  
|||  
Db 523 AGTGTGTTGGTGTCTATCCCGGAGTCTTACGTGAAGGACGCGCAAGGAGATGCCCTCCGCC 582  
|||  
QY 265 AAAGGTATACAGCTCCCAATGGATCAGTGGGAAGATCTGAGGGACAATATCAAGCTATA 324  
|||  
Db 583 AAAGGTATCAGTATGACTTTGGAGCAGTGGGAACCAATTTGCAATGTGTACCTGCAATA 642  
|||  
QY 325 GATAGGCCATCAAGGAGAAATGCGTGATCGGA 356  
|||  
Db 643 GAGGATGCCATCAAAAAGCTTGAAGATTTCAGA 674  
|||

RESULT 10  
US-10-425-114-21115  
; Sequence 21115, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 21115  
; LENGTH: 658  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3279-041-H1\_FLI  
US-10-425-114-21115

Query Match 11.8%; Score 74.8; DB 16; Length 658;  
Best Local Similarity 58.6%; Pred. No. 1.8e-12;  
Matches 130; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 135 CGCGACGACGATATCGTCGCGCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGAC 194  
|||  
Db 229 CGACCAAGGGGACCTTATTCTGTGCGCGCTTGTGCGAGCAAGAGGAGGTGACTCTGTCCGA 288  
|||  
QY 195 CTGGAACGGCAAGGTGCTGTCGACATCCGCGAGTCTTACGAGAAGGACGCGCAAGACCT 254  
|||  
Db 289 GTTCAAGGGCAGGTGCTGCTGTCCTCCAGAGTCTTACGTGAAGACGCGAAAGGAGAT 348  
|||  
QY 255 CCCCGCGCGCAAGGTATACAGCTCCCAATGGATCAGTGGGAAGATCTGAGGGACAATAT 314  
|||  
Db 349 GCCCTCCGCCAAAGGTATTAGTATGACAATGGAGCAGTGGGAAGCAATTTTGCATGTCTGT 408  
|||  
QY 315 CAAAGCTATAGTAGGCCATCAAGGAGAAATGCGTGATCGGA 356  
|||  
Db 409 ACCTGCGATAGAGCTGCCATCAAAAAGCTTGAAGATTTCAGA 450  
|||

RESULT 11  
US-10-425-115-20439  
; Sequence 20439, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 20439  
; LENGTH: 658  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_118642C.1  
US-10-425-115-20439

Query Match 11.8%; Score 74.8; DB 18; Length 658;  
Best Local Similarity 58.6%; Pred. No. 1.8e-12;  
Matches 130; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 135 CGCGACGACGATATCGTCGCGCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGAC 194  
|||  
Db 229 CGACCAAGGGGACCTTATTCTGTGCGCGCTTGTGCGAGCAAGAGGAGGTGACTCTGTCCGA 288  
|||  
QY 195 CTGGAACGGCAAGGTGCTGTCGACATCCGCGAGTCTTACGAGAAGGACGCGCAAGACCT 254  
|||  
Db 289 GTTCAAGGGCAGGTGCTGCTGTCCTCCAGAGTCTTACGTGAAGACGCGAAAGGAGAT 348  
|||  
QY 255 CCCCGCGCGCAAGGTATACAGCTCCCAATGGATCAGTGGGAAGATCTGAGGGACAATAT 314  
|||  
Db 349 GCCCTCCGCCAAAGGTATTAGTATGACAATGGAGCAGTGGGAAGCAATTTTGCATGTCTGT 408  
|||  
QY 315 CAAAGCTATAGTAGGCCATCAAGGAGAAATGCGTGATCGGA 356  
|||  
Db 409 ACCTGCGATAGAGCTGCCATCAAAAAGCTTGAAGATTTCAGA 450  
|||

## RESULT 12

US-10-767-701-29085  
; Sequence 29085, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 29085  
; LENGTH: 471  
; TYPE: DNA  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 8548715  
US-10-767-701-29085

Query Match 11.6%; Score 73.2; DB 17; Length 471;  
Best Local Similarity 58.1%; Pred. No. 4.8e-12;  
Matches 129; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 135 CGCGACGACGATATCGTCGCGCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGAC 194  
|||  
Db 37 CGACCAAGGAGACTTTATCTGTGCGCGCTGTGCGAACAGAGGAGGTGACTCTGTCCGA 96  
|||  
QY 195 CTGGAACGGCAAGGTGCTGTCGACATCCGCGAGTCTTACGAGAAGGACGCGCAAGACCT 254  
|||  
Db 97 GTTCAAGGGCAGGTGCTGCTGTCCTCCAGAGTCTTACGTGAAGGATGCGCAAGGAGAT 156  
|||  
QY 255 CCCCGCGCGCAAGGTATACAGCTCCCAATGGATCAGTGGGAAGATCTGAGGGACAATAT 314  
|||  
Db 157 GCCCTCGGCCAAAGGTATTAGTATGACGATGACGAGTGGAGAGCAATTTTGCATGTCTGT 216  
|||

Qy 315 CAAAGCTATAGATGAGGCCATCAAGGAGAAATGCGTGATCGGA 356  
 Db 217 ACCTGCAATAGATGATCCCATAAAGATTGAGATTGAGA 258

RESULT 13  
 US-10-424-599-127795/c  
 ; Sequence 127795, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J  
 ; APPLICANT: Kovalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 127795  
 ; LENGTH: 497  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_86402C.1  
 US-10-424-599-127795

Query Match 10.7%; Score 67.4; DB 16; Length 497;  
 Best Local Similarity 57.3%; Pred. No. 3.2e-10;  
 Matches 122; Conservative 0; Mismatches 91; Indels 0; Gaps 0;  
 Qy 139 GACGACGATATCGTCGCGCCAGATATCGAAGAACAGGAGGTGGCGGTGCGGACCTGG 198  
 Db 474 GAAGGCGATCTCATCTCGAGCTTTCAGATTAAGAGAGGGTGACGATTCAGGATTC 415  
 Qy 199 AACGGCAAGTCTGTCGCGCATCCGCGAGTCTTACGAGAGGAGCGGCAAGACCCCTCCC 258  
 Db 414 AGAGGGAACATTTGGTCTCCATTCGGGAGTATTATAAAGAGTGGCAAGGAGCTTCT 355  
 Qy 259 GGCGGCAAGGTATACAGTCCCAATGATCGATGAGGAGATCTAGGAGGCAATATCAAA 318  
 Db 354 ACTTCCAAAGGAATAAGTTTGACAGAGAGCGAGTGTCAACCTTTAAGAAAAATGTGCCT 295  
 Qy 319 GCTATAGATGAGGCCATCAAGGAGAAATGCGTGATCGGAGCCCAATCTCTTTG 369  
 Db 294 GCCATAGAAAAGCCATTAAAGAAAATGGAGTCA 262

RESULT 14  
 US-10-425-114-20666  
 ; Sequence 20666, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(5313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 20666  
 ; LENGTH: 794  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: LIB3242-699-H3\_FLI  
 US-10-425-114-20666

Query Match 10.5%; Score 66.2; DB 16; Length 794;  
 Best Local Similarity 55.4%; Pred. No. 9.5e-10;  
 Matches 128; Conservative 0; Mismatches 103; Indels 0; Gaps 0;  
 Qy 139 GACGACGATATCGTCGCGCCAGATATCGAAGAACAGGAGGTGGCGGTGCGGACCTGG 198  
 Db 373 GAAGGCGATCTCATCTCGAGCTTTCAGATTAAGAGAGGGTGACGATTCAGGATTC 432  
 Qy 199 AACGGCAAGTCTGTCGCGCATCCGCGAGTCTTACGAGAGGAGCGGCAAGACCCCTCCC 258  
 Db 433 AGAGGGAACATTTGGTCTCCATTCGGGAGTATTATAAAGAGTGGCAAGGAACTTCT 492  
 Qy 259 GGCGGCAAGGTATACAGTCCCAATGATCGATGAGGAGATCTAGGAGGCAATATCAAA 318  
 Db 493 ACTTCCAAAGGAATAAGTTTGACAGAGGAGCACTGGTCAGCCTTTAAGAAAATGTGCCT 552  
 Qy 319 GCTATAGATGAGGCCATCAAGGAGAAATGCGTGATCGGAGCCCAATCTCTTTG 369  
 Db 553 GCCATAGAAAAGCCATTAAAGAAAATGGAGTCAAGTTGACACATGGCTTTG 603

RESULT 15  
 US-10-424-599-4356  
 ; Sequence 4356, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J  
 ; APPLICANT: Kovalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 4356  
 ; LENGTH: 911  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_10393C.1  
 US-10-424-599-4356

Query Match 10.5%; Score 66.2; DB 16; Length 911;  
 Best Local Similarity 55.4%; Pred. No. 1e-09;  
 Matches 128; Conservative 0; Mismatches 103; Indels 0; Gaps 0;  
 Qy 139 GACGACGATATCGTCGCGCCAGATATCGAAGAACAGGAGGTGGCGGTGCGGACCTGG 198  
 Db 443 GAAGGCGATCTCATCTCGAGCTTTCAGATTAAGAGAGGGTGACGATTCAGGATTC 502  
 Qy 199 AACGGCAAGTCTGTCGCGCATCCGCGAGTCTTACGAGAGGAGCGGCAAGACCCCTCCC 258  
 Db 503 AGAGGGAACATTTGGTCTCCATTCGGGAGTATTATAAAGAGTGGCAAGGAACTTCT 562  
 Qy 259 GGCGGCAAGGTATACAGTCCCAATGATCGATGAGGAGATCTAGGAGGCAATATCAAA 318  
 Db 563 ACTTCCAAAGGAATAAGTTTGACAGAGGAGCACTGGTCAGCCTTTAAGAAAATGTGCCT 622  
 Qy 319 GCTATAGATGAGGCCATCAAGGAGAAATGCGTGATCGGAGCCCAATCTCTTTG 369  
 Db 623 GCCATAGAAAAGCCATTAAAGAAAATGGAGTCAAGTTGACACATGGCTTTG 673

Search completed: January 11, 2005, 23:15:46  
 Job time : 486 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 11, 2005, 19:53:20 ; Search time 2756 Seconds  
(without alignments)  
8356.281 Million cell updates/sec

Title: US-10-629-953-3  
Perfect score: 632  
Sequence: 1 ctctccagactccacag.....taaaaaaaaaaaaaaaaaaaaaa 632

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_ges1:\*  
9: gb\_ges2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	609.2	96.4	697	6	CB684473
2	609.2	96.4	784	6	CB684474
3	608.8	96.3	693	1	AU082888
4	501.4	79.3	539	1	AU092741
5	453.2	71.7	530	7	D15350
6	427.8	67.7	555	4	BI795760
7	387.6	61.3	401	1	AU068340
8	321.2	50.8	587	9	CL739643
9	277.6	43.9	643	8	AQ158085
10	268	42.4	396	6	C27852
11	262.2	41.5	329	2	BE607421
12	254.4	40.3	567	8	AQ795628
13	241.6	38.2	631	6	CA151445
14	240	38.0	443	6	C28075
15	240	38.0	523	6	CA116904
16	240	38.0	649	6	CA189554
17	239.6	37.9	292	1	AU183581
18	236.8	37.5	440	7	CF760736
19	236.8	37.5	635	6	CA230818
20	234.2	37.1	855	6	CA129813
21	234	37.0	637	6	CA259070
22	231.8	36.7	280	5	EX899341
23	224	35.4	564	6	CB863080
24	222.4	35.2	606	6	CA189705

## ALIGNMENTS

RESULT 1  
LOCUS CB684473  
DEFINITION OSJNEF14A12.f OSJNEf Oryza sativa (japonica cultivar-group) cDNA  
697 bp mRNA linear EST 09-APR-2003  
clone OSJNEF14A12 5', mRNA sequence.

ACCESSION CB684473  
VERSION CB684473.1  
KEYWORDS EST.  
SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.  
1 (bases 1 to 697)  
Jantauriyarat C., Lu G., Gowda M., Hatfield J., Zhou B., Mazur E., Kudrna D., Dean R., Soderlund C., Wing R. and Wang G.

Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe grisea  
Unpublished (2003)  
Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA  
Tel: 520 626 3967  
Fax: 520 621 9288  
Email: http://genome.arizona.edu  
PCR Primers  
FORWARD: gta aac cga cgg cca gtg  
BACKWARD: gga aac agc tat gac cat g  
Plate: 14 row: A column: 12  
Seq primer: gta aac cga cgg cca gtg.

FEATURES  
source

1. 697  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="OSJNEF14A12"  
/tissue\_type="Leaf"  
/dev\_stage="3 week"  
/lab\_host="DH10B"  
/clone\_lib="OSJNEf"  
/notes="Vector: pBluescript II KS +; Site\_1: EcoRI; Site\_2: XhoI; Uninfected Control"

ORIGIN

25 222.2 35.2 470 6 CD913769  
26 221.6 35.1 403 2 B4244113  
27 220.8 34.9 532 5 BQ838824  
28 220.6 34.9 1072 7 CK212042  
29 220.4 34.9 460 6 CA710341  
30 218.8 34.6 579 5 BU998906  
31 218.8 34.6 618 6 CB860065  
32 218.8 34.6 618 6 CB860077  
33 217.2 34.4 619 5 BU998918  
34 211.8 33.5 288 1 AU068341  
35 209.6 33.2 568 6 CB250145  
36 209.6 33.2 596 6 CD568912  
37 206.8 32.7 586 6 CD568913  
38 203.8 32.2 618 1 AI734589  
39 203.6 32.2 603 1 AI691445  
40 202.6 32.1 566 2 BE186543  
41 201.2 31.8 516 6 CF019517  
42 201.2 31.8 588 7 CF625939  
43 201.2 31.8 588 7 CF627159  
44 199 31.5 652 3 AY112536  
45 197.6 31.3 553 9 CL715353

```
Query Match      96.4%; Score 609.2; DB 6; Length 697;
Best Local Similarity 99.5%; Pred. No. 4.4e-155;
Matches 611; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTCCTCAGACTCCACAGTCCAGAGGGAAGTGTGAGAAGATGTGGCGGAAGGG 60
   |||||
Db 10 CTCCTCTAGACTCCACAGTCCAGAGGGAAGTGTGAGAAGATGTGGCGGAAGGG 69
   |||||
QY 61 AACAAAGCGTTCGGCGGCGGAGCGCGGCGCAAGCGCGTGCCTCGCGGGGACGAC 120
   |||||
Db 70 AACAAAGCGTTCGGCGGCGGAGCGCGGCGCAAGCGCGTGCCTCGCGGGGACGAC 129
   |||||
QY 121 GGCCCTCCGAGAGCGCGACGACGATATCGTGTGCGCCAGATATCGAAGAACAGAGG 180
   |||||
Db 130 GGCCCTCCGAGAGCGCGACGACGATATCGTGTGCGCCAGATATCGAAGAACAGAGG 189
   |||||
QY 181 GTGGCGGTGGGACCTGGAAAGCGGAGGTCGTGTGACATCCGCGAGTCTACGAGAAG 240
   |||||
Db 190 GTGGCGGTGGGACCTGGAAAGCGGAGGTCGTGTGACATCCGCGAGTCTACGAGAAG 249
   |||||
QY 241 GACGGCAAGACCTCCCGCGCGCAAGGTATACAGTCCCAATGGATCAGTGGAAAGATA 300
   |||||
Db 250 GACGGCAAGACCTCCCGCGCGCAAGGTATACAGTCCCGATGGATCAGTGGAAAGATA 309
   |||||
QY 301 CTGAGGGAACAATATCAAGCTATAGATGAGGCCATCAAGGAGAAATCGTGATCGAGCCC 360
   |||||
Db 310 CTGAGGGAACAATATCAAGCTATAGATGAGGCCATCAAGGAGAAATCGTGATCGAGCCC 369
   |||||
QY 361 ATTCTCTGTGATGCAAGTAGACTAAGCTACGTGTCTTTTATGACTCCGAGGAATAT 420
   |||||
Db 370 ATTCTCTGTGATGCAAGTAGACTAAGCTACGTGTCTTTTATGACTCCGAGGAATAT 429
   |||||
QY 421 TGCACTTTTGGTATGTAATCTTTATACCTGTCTAGTAATATCAAGTAATGTTTAGT 480
   |||||
Db 430 TGCACTTTTGGTATGTAATCTTTATACCTGTCTAGTAATATCAAGTAATGTTTAGT 499
   |||||
QY 481 TTGGTCATGCGCGGAAATGTGGCTGTCTGTGTTATGTGCCAAAGAGCAGCTATGTGC 540
   |||||
Db 490 TTGGTCATGCGCGGAAATGTGGCTGTCTGTGTTATGTGCCAAAGAGCAGCTATGTGC 549
   |||||
QY 541 CTAATGTGCTCTCTCTCAATTTGTACGTTGGGTGAAGAGGTCCATCTGTTAATGGTGT 600
   |||||
Db 550 CTAATGTGCTCTCTCTCAATTTGTACGTTGGGTGAAGAGGTCCATCTGTTAATGGTGT 609
   |||||
QY 601 GGTTCCTGCGTGTA 614
   |||||
Db 610 GGTTCCTGCGTGTA 623

RESULT 2
CB684474/c
LOCUS
DEFINITION
  Oryza sativa (japonica cultivar-group) cDNA
  clone OSJNEF14A12 3', mRNA sequence.
ACCESSION
  CB684474
  CB684474.1 GI:29688199
KEYWORDS
  EST.
SOURCE
  Oryza sativa (japonica cultivar-group)
  Oryza sativa (japonica cultivar-group)
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
  Ehrhartoideae; Oryzeae; Oryza.
  1 (bases 1 to 784)
REFERENCE
  Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
  Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  Large-scale identification of ESTs involved in the interaction
  between rice and Magnaporthe grisea
  Unpublished (2003)
  Contact: Rod Wing
  Arizona Genomics Institute
  University of Arizona
  Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
  85721-0088, USA
```

```
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
  FORWARD: gta aaa cga cgg cca gta g
  BACKWARD: gga aac agc tat gac cat g
  Plate: 14 row: A column: 12
  Seq primer: gga aac agc tat gac cat g.
FEATURES
  source
    1..784
    /organism="Oryza sativa (japonica cultivar-group)"
    /mol_type="mRNA"
    /cultivar="Nipponbare"
    /db_xref="taxon:39947"
    /clone="OSJNEF14A12"
    /tissue_type="Leaf"
    /dev_stage="3 week"
    /lab_host="DH10B"
    /clone_lib="OSJNEF"
    /notes="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
    XhoI; Uninfected Control"
ORIGIN
  Query Match      96.4%; Score 609.2; DB 6; Length 784;
  Best Local Similarity 99.5%; Pred. No. 4.5e-155;
  Matches 611; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTCCTCAGACTCCACAGTCCAGAGGGAAGTGTGAGAAGATGTGGCGGAAGGG 60
   |||||
Db 685 CTCCTCAGACTCCACAGTCCAGAGGGAAGTGTGAGAAGATGTGGCGGAAGGG 626
   |||||
QY 61 AACAAAGCGTTCGGCGGCGCGGAGCGCGCGCAAGCGCGTGCCTCGCGGGGACGAC 120
   |||||
Db 625 AACAAAGCGTTCGGCGGCGCGGAGCGCGCGCAAGCGCGTGCCTCGCGGGGACGAC 566
   |||||
QY 121 GGCCCTCCGAGAGCGCGACGACGATATCGTGTGCGCCAGATATCGAAGAACAGAGG 180
   |||||
Db 565 GGCCCTCCGAGAGCGCGACGACGATATCGTGTGCGCCAGATATCGAAGAACAGAGG 506
   |||||
QY 181 GTGGCGGTGGGACCTGGAAAGCGGAGGTCGTGTGACATCCGCGAGTCTACGAGAAG 240
   |||||
Db 505 GTGGCGGTGGGACCTGGAAAGCGGAGGTCGTGTGACATCCGCGAGTCTACGAGAAG 446
   |||||
QY 241 GACGGCAAGACCTCCCGCGCGCAAGGTATACAGTCCCAATGGATCAGTGGAAAGATA 300
   |||||
Db 445 GACGGCAAGACCTCCCGCGCGCAAGGTATACAGTCCCGATGGATCAGTGGAAAGATA 386
   |||||
QY 301 CTGAGGGAACAATATCAAGCTATAGATGAGGCCATCAAGGAGAAATGCGTGATCGGAGCCC 360
   |||||
Db 385 CTGAGGGAACAATATCAAGCTATAGATGAGGCCATCAAGGAGAAATGCGTGATCGGAGCCC 326
   |||||
QY 361 ATTCTCTGTGATGCAAGTAGACTAAGCTACGTGTCTTTTATGACTCCGAGGAATAT 420
   |||||
Db 325 ATTCTCTGTGATGCAAGTAGACTAAGCTACGTGTCTTTTATGACTCCGAGGAATAT 266
   |||||
QY 421 TGCACTTTTGGTATGTAATCTTTATACCTGTCTAGTAATATCAAGTAATGTTTAGT 480
   |||||
Db 265 TGCACTTTTGGTATGTAATCTTTATACCTGTCTAGTAATATCAAGTAATGTTTAGT 206
   |||||
QY 481 TTGGTCATGCGCGGAAATGTGGCTGTCTGTTATGTGCCAAAGAGCAGCTATGTGC 540
   |||||
Db 205 TTGGTCATGCGCGGAAATGTGGCTGTCTGTTATGTGCCAAAGAGCAGCTATGTGC 146
   |||||
QY 541 CTAATGTGCTCTCTCTCAATTTGTACGTTGGGTGAAGAGGTCCATCTGTTAATGGTGT 600
   |||||
Db 145 CTAATGTGCTCTCTCTCAATTTGTACGTTGGGTGAAGAGGTCCATCTGTTAATGGTGT 86
   |||||
QY 601 GGTTCCTGCGTGTA 614
   |||||
Db 85 GGTTCCTGCGTGTA 72

RESULT 3
```

```
AU082888
LOCUS
DEFINITION AU082888 Rice callus Oryza sativa (japonica cultivar-group) cDNA
            clone C53889, mRNA sequence.
ACCESSION AU082888.1 GI:7212585
VERSION
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Oryza.
1 (bases 1 to 693)
Sasaki, T. and Yamamoto, K.
Rice cDNA from callus (2000)
Unpublished (2000)
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
PROJECT = 'RGP'.
Location/Qualifiers
1..693
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="C53889"
/clone_lib="Rice callus"
/notes="vector: pBluescript II SK+; Site 1: SalI; Site 2:
NotI; cDNA prepared from rice callus mRNAs by using
oligo(dT) as a primer and ligating to the SalI-NotI site
of pBluescript II SK+ phagemid."
FEATURES
source
Query Match 96.3%; Score 608.8; DB 1; Length 693;
Best Local Similarity 99.3%; Pred. No. 5.7e-155;
Matches 610; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
ORIGIN
1 CTCTTCAAGACTCCACAGTCCAGAGGGAGAAAGTGTGAGAAAGTGGCGGAAGGG 60
17 CTCTTCAAGACTCCACAGTCCAGAGGGAGAAAGTGTGAGAAAGTGGCGGAAGGG 76
61 AACAGCGGTTCGGCGGGCGGAGCGCGCGCCCAAGCGCGCTGCGCGGGGAGCAC 120
77 AACAGCGGTTCGGCGGGCGGAGCGCGCGCCCAAGCGCGCTGCGCGGGGAGCAC 136
121 GGGCCCTCCGAGAGCGCGGACGATATCGTGTGCGCCAGATATCGAAGAACAGGAGG 180
137 GGGCCCTCCGAGAGCGCGGACGATATCGTGTGCGCCAGATATCGAAGAACAGGAGG 196
181 GTGGCGGTGGGACCTGGAACGGCAAGGTGTGTGTGACATCCCGAGTCTACGAGAAG 240
197 GTGGCGGTGGGACCTGGAACGGCAAGGTGTGTGTGACATCCCGAGTCTACGAGAAG 256
241 GACGGCAAGACCTCCCGCGCGCAAGAGTATACAGTCCCAATGGATCAGTGAAGATA 300
257 GACGGCAAGACCTCCCGCGCGCAAGAGTATACAGTCCCGAGTCTACGAGAAGATA 316
301 CTGAGGGACAAATCAAGAGCTATAGATGAGGCCATCAAGGAGATGCGTATCGGAGCCC 360
317 CTGAGGGACAAATCAAGAGCTATAGATGAGGCCATCAAGGAGATGCGTATCGGAGCCC 376
361 ATTCTTTGTGATGCAAGTAGACTAAGCTTAGCTGTCTTTTATGATCTCGAGGAATAT 420
377 ATTCTTTGTGATGCAAGTAGACTAAGCTTAGCTGTCTTTTATGATCTCGAGGAATAT 436
421 TGCATTTTGGTAGTGAATCTTTTATACCTGTCTAGTAATATCAAGTAATGTTTAGT 480
437 TGCATTTTGGTAGTGAATCTTTTATACCTGTCTAGTAATATCAAGTAATGTTTAGT 496
```

```
Qy 481 TTGGTCATGCGCGGAAAATGTGGCTTGTCTGGTATTGTGCCAAAGAGCAGCTATGCG 540
Db 497 TTGGTCATGCGCGGAAAATGTGGCTTGTCTGGTATTGTGCCAAAGAGCAGCTATGCG 556
Qy 541 CTAATGTGCTCTCTCTCATTTTGTACGTTGGTGAAGAGGTCCATCTGTTAATGTTGGTT 600
Db 557 CTAATGTGCTCTCTCTCATTTTGTACGTTGGTGAAGAGGTCCATCTGTTAATGTTGGTT 616
Qy 601 GGTTCCTGCGGTGA 614
Db 617 GGTTCCTGCGGTGA 630
RESULT 4
LOCUS
DEFINITION AU092741 Rice callus Oryza sativa (japonica cultivar-group) cDNA
            clone C53228, mRNA sequence.
ACCESSION AU092741
VERSION
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Oryza.
1 (bases 1 to 539)
Sasaki, T. and Yamamoto, K.
Rice cDNA from callus (2000)
Unpublished (2000)
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
PROJECT = 'RGP'.
Location/Qualifiers
1..539
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="C53228"
/clone_lib="Rice callus"
/notes="vector: pBluescript II SK+; Site 1: SalI; Site 2:
NotI; cDNA prepared from rice callus mRNAs by using
oligo(dT) as a primer and ligating to the SalI-NotI site
of pBluescript II SK+ phagemid."
FEATURES
source
Query Match 79.3%; Score 501.4; DB 1; Length 539;
Best Local Similarity 96.7%; Pred. No. 1.1e-125;
Matches 523; Conservative 0; Mismatches 16; Indels 2; Gaps 1;
ORIGIN
81 CGGCGAGCCGCGCCCAAGCGCGCTGCGCGCGGAGACGAGCGGCCCTCCGAGAGCGCGA 140
Db 1 CGGCGAGCCGCGCGCCCAAGCGCGCTGCGCGCGGAGACGAGCGGCCCTCCGAGAGCGCGA 60
Qy 141 CGACGATATCGTGTGCTGCCAGATATCGAAGAACAGAGGGTGGCGGTGCGGAGCTGGA 200
Db 61 CGACGATATCGTGTGCTGCCAGATATCGAAGAACAGAGGGTGGCGGTGCGGAGCTGGA 118
Qy 201 CGGCAAGGTGTGTGTGCTGACATCCCGAGTCTACGAGAGACGCGAAGACCTCCCGG 260
Db 119 CGGCAAGGTGTGTGTGCTGACATCCCGAGTCTACGAGAGACGCGAAGACCTCCCGG 178
Qy 261 CCGCAAGGTATACAGCTCCCAATGGATCAGTGAAGATCTAGAGGACAAATATCAAGC 320
Db 179 CCGCAAGGTATACAGCTCCCGATGGATCAGTGAAGATCTAGAGGACAAATATCAAGC 238
```



QY 321 TATAGATAGGCCCATCAGGAGAAATGCGTATCGAGCCCATCTCTTGTGATGCAAGTA 380  
 |||||  
 Db 239 TATAGATAGGCCCATCAGGAGAAATGCGTATCGAGCCCATCTCTTGTGATGCAAGTA 298  
 |||||  
 QY 381 GACTAGCCTACGCTCTGCTTTTATGACTCCGAGGATATGACATTTTGGTATGTTAAT 440  
 |||||  
 Db 299 GACTAGCCTACGCTCTGCTTTTATGACTCCGAGGATATGACATTTTGGTATGTTAAT 358  
 |||||  
 QY 441 CTTTATTACCTCTCTAGTAAATCAAGTAATGTTTGTGATGCTCATGCGCGAAATG 500  
 |||||  
 Db 359 CTTTATTACCTCTCTAGTAAATCAAGTAATGTTTGTGATGCTCATGCGCGAAATG 418  
 |||||  
 QY 501 TGCCCTTGCTGATTTGTCCTAAAGAGCAGCTATGTCCTAATGCTCTCTCATTT 560  
 |||||  
 Db 419 TGCCCTTGCTGATTTGTCCTAAAGAGCAGCTATGTCCTAATGCTCTCTCATTT 478  
 |||||  
 QY 561 TGACGTTGGTAAGAGGTCCTATCTGTTAATGGTTGGTTGTTTCTGCGTGTAATAAAA 620  
 |||||  
 Db 479 TGACGTTGGTAAGAGGTCCTATCTGTTAATGGTTGGTTGTTTCTGCGTGTAATAAAA 538  
 |||||  
 QY 621 A 621  
 |||||  
 Db 539 A 539  
 |||||

RESULT 5  
 D15350  
 LOCUS  
 DEFINITION  
 ACCESION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

D15350 530 bp mRNA linear EST 24-FEB-2003  
 R100499A Rice callus Oryza sativa (japonica cultivar-group) cDNA  
 clone C0499, mRNA sequence.  
 D15350 C98059  
 D15350.2 GI:15072274  
 EST.  
 Oryza sativa (japonica cultivar-group)  
 Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 COMMENT

1 (bases 1 to 530)  
 Sasaki,T., Song,J., Koga-Ban,Y., Matsui,E., Fang,F., Higo,H.,  
 Nagabaki,H., Hori,M., Miya,M. and Murayama-Kayano,E.  
 Toward cataloging all rice genes: large-scale sequencing of  
 randomly chosen rice cDNAs from a callus cDNA library  
 Plant J. 6 (4), 615-624 (1994)  
 95078950  
 7987417  
 On May 17, 1993 this sequence version replaced gi:286543  
 gi:3760805.  
 Contact: Takuji Sasaki  
 National Institute of Agrobiological Resources  
 Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
 305-8602, Japan  
 Tel: 81-298-38-7441  
 Fax: 81-298-38-7468  
 Email: tsasaki@nri.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
 PROJECT = 'RGP'  
 Seq primer: oligo(dT).  
 Location/Qualifiers

FEATURES  
 source  
 1..530  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="C0499"  
 /clone\_lib="Rice callus"  
 /notes="Vector: pBluescript II SK-; Site 1: SalI; Site 2:  
 NotI; cDNA prepared from rice callus mRNAs by using  
 oligo(dT) as a primer and ligating to the SalI-NotI site  
 of pBluescript II SK+ phagemid."  
 ORIGIN

Query Match 71.7%; Score 453.2; DB 7; Length 530;  
 Best Local Similarity 97.5%; Pred. No. 1.6e-112;

Matches 499; Conservative 0; Mismatches 9; Indels 4; Gaps 4;  
 QY 1 CTCTCTCAAGACTCCACAAGTCCAGAGGAGAAAGTGTGAGAAATGTGCGGAGGG 60  
 |||||  
 Db 21 CTCTCTCAAGACTCCACAAGTCCAGAGGAGAAAGTGTGAGAAATGTGCGGAGGG 80  
 |||||  
 QY 61 AACAAAGGGTTCCG - CGGCGCGCGGAGCCGCGGCGCAAGCGCGCTGCGCCCGGGACGA 119  
 |||||  
 Db 81 AACAAAGGGTTCCGAGCGCGCGCGCGCAAGCGCGCTGCGCCCGGGACGA 140  
 |||||  
 QY 120 CGGGCCCTCCGAGAGCGCGCGAGCCAGCATATCGTCGCCAGATATCGAAGACAGGAG 179  
 |||||  
 Db 141 CGGGCCCTCCGAGAGCGCGCGAGCCAGCATATCGTCGNGCCAGATATCGAAGACAGGAG 200  
 |||||  
 QY 180 GGTGGCGGTGCGGACCTCGGAAACGGCAAGGTGCTCGTCGACATCCGCGAGTTCTACGAGAA 239  
 |||||  
 Db 201 GGTGGCGGTGCGGACCTCGGAAACGGCAAGGTGCTCGTCGACATCCGCGAGTTCTACGAGAA 260  
 |||||  
 QY 240 GGAACGCAAGACCTCCCGCGCGCAAGGTATACAGCTCCCAATGGATCAGTGGAAGAT 299  
 |||||  
 Db 261 GGACGCGAAGACCTCCCGCGCGCAAGGTATACAGCTCCCAATGGATCAGTGGAAGAT 320  
 |||||  
 QY 300 ACTGAGGACAAATATCAAGCTATAGATGAGGCCATCAAGAGAAATCGGTATCGGAGCC 359  
 |||||  
 Db 321 ACTGAGGACAAATATCAAGCTATAGATGAGGCCATCAAGAGAAATCGGTATCGGAGCC 379  
 |||||  
 QY 360 CATTCCTCTGTGATGCAAGTAGACTAAGCCCTAGCTCTGTCTTTTATGACTCCGAGGAAATA 419  
 |||||  
 Db 380 CATTCCTCTGTGATGCAAGTAGACTAAGCCCTAGCTCTGTCTTTTATGACTCCGAGGAAATA 439  
 |||||  
 QY 420 TTGCACCTTTTGGTATGTTAATCTTTTATACCTGTCTAGTAAATATCAAGTAATGTTTTAG 479  
 |||||  
 Db 440 TTGCACCTTTTGGTATGTTAATCTTTTATACCTGTCTAGTAAATATCAAGTAATGTTTTAG 498  
 |||||  
 QY 480 TTGGTCTATCGCGCGG - AAAATGTGGCTTGTGTC 510  
 |||||  
 Db 499 TTTGGTCAAGCGCGGAAAAAATGTGGCTTGTGTC 530  
 |||||

RESULT 6  
 BI795760  
 LOCUS  
 DEFINITION  
 ACCESION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

BI795760 555 bp mRNA linear EST 02-OCT-2001  
 H030F04 Endosperm library from Oryza sativa (10 days after  
 anthesis) Oryza sativa cDNA clone H030F04, mRNA sequence.  
 BI795760  
 BI795760.1 GI:15847484  
 EST.  
 Oryza sativa  
 Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.  
 1 (bases 1 to 555).  
 Dong,H.T., Li,D.B., Zhuang,X.F., Dai,C.G., Sun,L.X., Pei,Y.X.,  
 Wu,H.F., Jiang,Y.X., Yu,F.C., Gao,Q.K. and Lou,Y.C.  
 A Gene Expression Screen in Oryza sativa  
 Unpublished (2001)  
 Contact: Haitao Dong, Debao Li  
 Bioinformatics and Gene Network Research Group  
 Zhejiang University  
 Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China  
 Tel: 0086-571-86892051  
 Fax: 0086-571-86961525  
 Email: webmaster@estarray.org, URL: http://www.estarray.org  
 Seq primer: M13 forward primer.  
 Location/Qualifiers

FEATURES  
 source  
 1..555  
 /organism="Oryza sativa"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4530"  
 /clone="H030F04"  
 /tissue\_type="Endosperm"  
 /dev\_stage="10 days after anthesis"





Email: <http://genome.arizona.edu>  
 PCR Primers  
 FORWARD: TAA TAC GAC TCA CTA TAG GG  
 BACKWARD: CAC TCA TTA GGC ACC CCA  
 Insert length: 161 Std Error: 0.00  
 Plate: 0074 row: L column: 07  
 Seq primer: TAA TAC GAC TCA CTA TAG GG  
 Class: BAC ends.

# FEATURES

source Location/Qualifiers

1..587  
 /organism="Oryza rufipogon"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:4529"  
 /clone="OR\_BBa0074L07"  
 /tissue\_type="young leaves"  
 /lab\_host="DH10B-T1 phage resistant"  
 /clone\_lib="OR\_BBa"  
 /notes="Vector: pAGIBAC1; Site\_1: HindIII; Site\_2: HindIII"

# ORIGIN

Query Match 50.8%; Score 321.2; DB 9; Length 587;  
 Best Local Similarity 99.1%; Pred. No. 1.7e-76;  
 Matches 323; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 289 CAGTGGAGACTACTGAGGACAATATCAAGCTATAGATGAGGCCATCAAGGAGATGCG 348  
 236 CAGTGGAGACTACTGAGGACAATATCAAGCTATAGATGAGGCCATCAAGGAGATGCG 295  
 349 TGATCGAGGCCATTCCTTGTGATGCAAGTAGACTAGCTACGCTGCTCTTTATGAC 408  
 296 TGATCGAGGCCATTCCTTGTGATGCAAGTAGACTAGCTACGCTGCTCTTTATGAC 355  
 409 TCCGAGCAATATTGCACCTTTGGTATGTTATTTATACCTGCTAGTAAATCAAG 468  
 356 TCCGAGCAATATTGCACCTTTGGTATGTTATTTATACCTGCTAGTAAATCAAG 415  
 469 TAATGTTTATGTTGGTCATCGCGGAAATATGTCCTGCTGTTATGTCGCAAGAA 528  
 416 TAATGTTTATGTTGGTCATCGCGGAAATATGTCCTGCTGTTATGTCGCAAGAA 475  
 529 GCAGCTATGTCCTTAATGTCCTCTCTCATTTGTTGTTGTTGTTGTTGTTGTTGTT 588  
 476 GCAGCTATGTCCTTAATGTCCTCTCTCATTTGTTGTTGTTGTTGTTGTTGTTGTT 535  
 589 TAATGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 614  
 536 TAATGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 561

# RESULT 9

AQ158085 643 bp DNA linear GSS 12-SEP-1998  
 LOCUS nbxb0010L09r CUGI Rice BAC Library Oryza sativa (japonica  
 DEFINITION cultivar-group) genomic clone nbxb0010L09r, genomic survey  
 sequence.

ACCESSION AQ158085

VERSION AQ158085.1 GI:3555110

# KEYWORDS

SOURCE Oryza sativa (japonica cultivar-group)

# ORGANISM

Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 1 (bases 1 to 643)

# REFERENCE

AUTHORS Wing,R.A. and Dean,R.A.  
 TITLE A BAC End Sequencing Framework to Sequence the Rice Genome  
 JOURNAL Unpublished (1998)

# COMMENT

Contact: Wing RA  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Tel: 864 656 7288  
 Fax: 864 656 4293

Email: [rwing@clemson.edu](mailto:rwing@clemson.edu)  
 Seq primer: GGAAACAGCTATGACCATG  
 Class: BAC ends  
 High quality sequence start: 25  
 High quality sequence stop: 517.  
 Location/Qualifiers

# FEATURES

source Location/Qualifiers  
 1..643  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="genomic DNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="nbxb0010L09r"  
 /tissue\_type="Leaf"  
 /lab\_host="E. coli DH10B"  
 /clone\_lib="CUGI Rice BAC Library"  
 /notes="Vector: pBelobAC11; Site\_1: HindIII; Site\_2:  
 HindIII; Rice is one of two most popular grains in the  
 world. Half of the world population especially those  
 inhabiting highly populated areas of the humid tropics  
 and subtropics, rely on rice as their primary source of  
 carbohydrate. Monocotyledonous rice is a diploid plant  
 (2n=24) with a haploid genome equivalent of 431 Mbp  
 (Arumuganathan and Earle, 1991). The relatively small  
 genome of rice, three times larger than that of  
 Arabidopsis, makes it suitable for genomic studies. In  
 order to facilitate positional cloning, physical mapping  
 and genome sequencing of rice, we have constructed a BAC  
 library from Oryza sativa, Nipponbare variety. The  
 library contains 36,864 clones with an average insert size  
 of 128.5 Kb providing 10.9 haploid genome equivalents. The  
 deep coverage allows the isolation a particular sequence  
 with a probability of 99.9 %. Two high density filters,  
 each containing 18,432 clones (doubly spotted), represent  
 the whole library for colony screening."

# ORIGIN

Query Match 43.9%; Score 277.6; DB 8; Length 643;  
 Best Local Similarity 98.3%; Pred. No. 1.3e-64;  
 Matches 291; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
 289 CAGTGGAGACTACTGAGGACAATATCAAGCTATAGATGAGGCCATCAAGGAGATGCG 348  
 348 CAGTGGAGACTACTGAGGACAATATCAAGCTATAGATGAGGCCATCAAGGAGATGCG 407  
 349 TGATCGAGGCCATTCCTTGTGATGCAAGTAGACTAGCTACGCTGCTCTTTATGAC 408  
 408 TGATCGAGGCCATTCCTTGTGATGCAAGTAGACTAGCTACGCTGCTCTTTATGAC 467  
 409 TCCGAGCAATATTGCACCTTTGGTATGTTATTTATACCTGCTAGTAAATCAAG 468  
 468 TCCGAGCAATATTGCCTTTT-GTATGGTAACTTTCATTACCTGCTAGTAAATCAAG 526  
 469 TAATGTTTATGTTGGTCATCGCGGAAATATGTCCTGCTGTTATGTCGCAAGAA 528  
 527 TAATGTTTATGTTGGTCATCGCGGAAATATGTCCTGCTGTTATGTCGCAAGAA 586  
 529 GCAGCTATGTCCTTAATGTCCTCTCTCATTTGTTGTTGTTGTTGTTGTTGTTGTT 584  
 587 GCAGCTATGTCCTTAATGTCCTCTCTCATTTATGTTGTTGTTGTTGTTGTTGTTGTT 642

# RESULT 10

# LOCUS

C27852 396 bp mRNA linear EST 04-APR-2002  
 DEFINITION C27852 Rice callus cDNA Oryza sativa (japonica cultivar-group) CDNA  
 clone C53228.1A, mRNA sequence.

ACCESSION C27852

VERSION C27852.1 GI:2311697

# KEYWORDS

SOURCE Oryza sativa (japonica cultivar-group)

# ORGANISM

Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzeae; Oryza.  
1 (bases 1 to 396)  
Yamamoto, K. and Sasaki, T.  
Rice cDNA from callus 1997  
Unpublished (1997)  
Contact: Takuji Sasaki  
National Institute of Agrobiological Resources  
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
305-8602, Japan  
Tel: 81-298-38-7441  
Fax: 81-298-38-7468  
Email: tsasaki@agr.affrc.go.jp. URL: <http://rgp.dna.affrc.go.jp/>  
PROJECT = 'RGP'.

FEATURES	Source	Location/Qualifiers
	1.396	
		/organism="Oryza sativa (japonica cultivar-group)"
		/mol_type="mRNA"
		/cultivar="Nipponbare"
		/db_xref="taxon:39947"
		/clone="C53228_1A"
		/tissue_type="callus"
		/dev_stage="callus"
		/clone_lib="Pice callus cDNA"

## ORIGIN

Query Match	42.4%;	Score 268;	DB 6;	Length 396;
Best Local Similarity	88.1%;	Pred. No. 5.1e-62;		
Matches 342;	Conservative 0;	Mismatches 37;	Indels 9;	Gaps 5;
QY	6	CCAAGACTCCACAAGTCAGAGGCGAGAAAGTGTGAGAAAGATGTGGCGGAGGGGAAACAA	65	
DB	1	CCAGACNCCACAAGTCCAGNGGGAGGAAGGTGTGAGAAAGATNTAGCNAGAAAGGGAACAA	60	
QY	66	GCGETTCGGCGGCGCGCGAGCGCGGCCAGCGCCGTGTCGCCCGGGGACGACGGGCC	125	
DB	61	GCGETTCNACAGCANCAAGCAAGCGCGCGCCAGATGCACCGNACCAACCGNGNACGACGGGCC	120	
QY	126	CTCCGAGAGCGCGACACGACGATATCGTCTGCCCGCAGATATCGAAGAACACGAGGGGTGGC	185	
DB	121	CTCCGAGAGCGCGACACGATATCGTCT--TNCCCAGATATCGAAGAACACGAGGGGTGGC	178	
QY	186	GGTGCAGACCTGGAAACGGCAAGGTCTGTCGCATATCCGGAGTTCTACAGAGAGGACGG	245	
DB	179	GGTGCAGACTGGAAACGGCAAGGTCTGTCGCATATCCGGAGTTCTACAGAGAGGACGG	238	
QY	246	CAAGACCTCTCCCGGCGGCAAGGTA-TACAGCTCCCAATGGATCAGTGGAAAGTACTGA	304	
DB	239	CAAGACCTCTCCCGGCGGCAAGGATTTACAGCTCCCGATGGATCAGTGGAAAGTACTTN	298	
QY	305	--GGGCAATATCAAGCTATAGATGAGGCCATCAAGGA-GAATGCGTGTATCGAGGCCCA	361	
DB	299	AGGGGCAATATCAAGCTATAGATGAGGCCATCAAGGAGGAATGCGTNATCGAGGCCCA	358	
QY	362	---TTCTCTGTGTGATGCAAGTAGACTAA	386	
DB	359	TTTTTCTTTGTGTGATGCAAGTAGACTTA	386	

RESULT 11  
BE607421

LOCUS	B5607421	329 bp	mRNA	linear	EST 23-AUG-2000
DEFINITION	OS202A06 OE <i>Oryza sativa</i> (indica cultivar-group) cDNA clone Rice Microarray ID699 similar to unknown, mRNA sequence.				
ACCESSION	B5607421				
VERSION	B5607421.1	GI:9886329			
KEYWORDS	EST.				
SOURCE	<i>Oryza sativa</i> (indica cultivar-group)				
ORGANISM	<i>Oryza sativa</i> (indica cultivar-group)				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; <i>Oryza</i> .				
REFERENCE	1 (bases 1 to 329)				
AUTHORS	Bohnert H.J., Bohnert C., Brazille S., Brooks J., Eaton M.,				

## REFERENCE AUTHORS

Perrea,H., Kawasaki,S., McCollough,A., Michalowski,C.B.,  
Palacio,C., Scara,G., Wheeler,M. and Zepeda,G.R.  
Functional Genomics of Plant Stress Tolerance  
Unpublished (2000)  
Contact: Michalowski,C.B.

CONFIDENTIAL

Contact: Michalowski, C.B.  
University of Arizona  
Bio Sciences West room 513, Tucson, AZ 85721, USA  
Tel: 520-621-7982  
Fax: 520-621-1697  
Email: cbm@u.arizona.edu

No hits found using blastx vs. genbank:(nr), cutoff:(e=0.1)  
Unknown, putative, similar, and hypothetical are not characterized  
with biological features.

**FEATURES**  
**source**

```

source
1. 329
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/strain="Pokkali"
/db_xref="taxon:39946"
/clone="Rice Microarray ID699"
/tissue_type="roots"
/dev_stage="1 week"
/clone_lib="OE"
/note="2-3 d 150mM NaCl"

```

## ORIGIN

Query Match	41.5%;	Score	262.2;	DB 2;	Length	329;			
Best Local Similarity	98.9%;	Pred. No.	1.9e-60;						
Matches	264;	Conservative	0;	Mismatches	3;	Indels	0;	Gaps	0;
Qy	348	GTGATCGAGGCCCATTCCTCTGTGTGATGCAAGTAGACTAAGCCTACGTCGTCTTTTATGA	407						
Db	9	GTGATCGAGGCCCATTCCTCTGTGTGATGCAAGTAGACTAAGCCTTCGTCTGTCTTTTATGA	68						
Qy	408	CTCCGAGGAATATTGCACCTTTTGGTAGTGGTAATCTTTATTACCCCTGCTAGTATATACAA	467						
Db	69	CTCCGAGGAATATTGCACCTTTTGGTAGTGGTAATCTTCACCTACCCCTGCTAGTATATACAA	128						
Qy	468	GTAATGTTTTAGTTTGGTTCATGCGCGGAAAATGTGGCCTTGTCTGGGTATTGTGCGCAAGA	527						
Db	129	GTAATGTTTTAGTTTGGTTCATGCGCGGAAAATGTGGCCTTGTCTGGGTATTGTGCGCAAGA	188						
Qy	528	AGCAGCTATGTGCCTAATGTGCTCTCCCTCATTTTGTACGTTGGGTAGAGGTCCTCATCTG	587						
Db	189	AGCAGCTATGTGCCTAATGTGCTCTCCCTCATTTTGTACGTTGGGTAGAGGTCCTCATCTG	248						
Qy	588	TTTAATGGTTGGTTGGTTTCTGCGTGTA	614						
Db	249	TTAATGGTTGGTTGGTTTCTGCGTGTA	275						

## RESULT 12

LOCUS	AQ795628	567 bp	DNA	linear	GSS 04-AUG-1999
DEFINITION	nbxb0056E24r CUGI Rice BAC Library Oryza sativa (japonica cultivar-group) genomic clone nbxb0056E24r, genomic survey sequence.				

sequence.  
ACCESSION AO795628

VERSION AQ795628.1 GI:5704013

**KEYWORDS** GSS.

**SOURCE** *Oryza sativa* (japonica cultivar-group)

ORGANISM  
*Oryza sativa* (japonica cultivar-group)  
Eukaryota; Varidiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Eriartoideae; Oryzaceae; *Oryza*.

REFERENCE 1 (bases 1 to 567)

**AUTHORS** Wing, R.A. and Dean, R.A.

**TITLE** A BAC End Sequencing Pr

**JOURNAL**  
**Unpublished (1998)**

COMMENT

Clemson University  
Clemson University  
Clemson University

Clemson University  
100 Jordan Hall, Clemson, SC 29634. USA

TOO MANY MATT, CLEMSON, SC 22033, USA



```
clone C53889_2A, mRNA sequence.
C28075
C28075.1 GI:2311920
EST.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartioideae; Oryzaceae; Oryza.
1 (bases 1 to 443)
Yamamoto,K. and Sasaki,T.
Rice cDNA from callus 1997
Unpublished (1997)
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@ab.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
PROJECT = 'RGP'.
FEATURES
    source
        1..443
            /organism="Oryza sativa (japonica cultivar-group)"
            /mol_type="mRNA"
            /cultivar="Nipponbare"
            /db_xref="taxon:39947"
            /clone="C53889_2A"
            /tissue_type="Callus"
            /dev_stage="callus"
            /clone_lib="Rice callus cDNA"
ORIGIN
Query Match      38.0%; Score 240; DB 6; Length 443;
Best Local Similarity 92.7%; Pred. No. 2.3e-54;
Matches 303; Conservative 0; Mismatches 17; Indels 7; Gaps 5;
QY 1 CTCTCTCAAGACTCCACAAGTCGAGGAGGAGAAAGTGTGAGAAGATGTGCGGAGGGG 60
DB 18 CTCTCTCAAGACTCCACAAGTCGAGGAGGAGAAAGTGTGAGAAGATGTGCGGAGGGG 77
QY 61 AACAGCGGTTCGCGCGCGCGGCGAGCGCGCCCAAGCGCGTTCGCGCGGCGGACGAC 120
DB 78 AACAGCGGTTCGCGCGCGCGGCGAGCGCGCGCCCAAGCGCGTTCGCGCGGCGGACGAC 137
QY 121 GGCCCTCCGAGAGCGCGGACGACGATATCGTTCGTCGCCAGATATCGAAGAACAGGAGG 180
DB 138 GGCCCTCCGAGAGCGCGGACGACGATATCGT-NTNCNCCAGATATCGAAGAACAGGAGG 196
QY 181 GTGCGGTGCGGACCTCGGAAACGGAAGGTGCTGTCGACATCCGCGAG-TTCTACGAGAA 239
DB 197 GTGCGGTGCGGACCTCGGAAACGGAAGGTGCTGTCGACATCCGCGNAGTTTNTACGAGAA 256
QY 240 GGAAGCGCAAGACCTCCCGCGCGC-AAAGGTATACAGC-TCCCNATGGATCAGTGG--- 294
DB 257 GGAAGCGCAAGACCTCCCGCGCGC-AAAGGTATACAGCTTCGCGATNGATCAGTTGGAA 316
QY 295 AAGATACTGAGGGACAATATCAAGCT 321
DB 317 GATACTTAGGGGNCATATCAAGCT 343
RESULT 15
LOCUS      523 bp mRNA linear EST 23-SEP-2003
DEFINITION SCAGLR1021A08.9 LR1 Saccharum officinarum cDNA clone SCAGLR1021A08
            5', mRNA sequence.
ACCESSION  C116904
VERSION    C116904.1 GI:34970211
KEYWORDS   EST.
SOURCE     Saccharum officinarum
ORGANISM   Saccharum officinarum; Streptophyta; Embryophyta; Tracheophyta;
            Eukaryota; Viridiplantae;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
1 (bases 1 to 523)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCSEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bccc.bccp.br
Plate: 021 row: A column: 08
Seq primer: T7 Promoter Primer.
Location/Qualifiers
    1..523
        /organism="Saccharum officinarum"
        /mol_type="mRNA"
        /db_xref="taxon:4547"
        /clone="SCAGLR1021A08"
        /lab_host="DH10B"
        /clone_lib="LR1"
        /notes="Organ: Leaf roll from field grown adult plants
        (large insert library); Vector: pSport1; Site 1: SalI;
        Site 2: NotI; An unidirectional cDNA library generated
        from [leaf roll from field grown adult plants (large
        insert library)]. cDNA was prepared from polyA+ mRNA
        using SuperScript Plasmid System Kit (Invitrogen). The
        double-strand cDNAs were fractionated in a sepharose
        CL-2B 40cm-columns and fragments sizing between 0.8 and
        1.5 Kb were directionally cloned into the vector. Details
        of each source of RNA and library construction can be
        obtained at http://sucsest.lad.ic.unicamp.br/public"
ORIGIN
Query Match      38.0%; Score 240; DB 6; Length 523;
Best Local Similarity 84.4%; Pred. No. 2.3e-54;
Matches 270; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 40 AGAAGATGTGCGGAGGAGGAAACAAGCGTTCGCGCGCGCGCGCGGCGGCGGCGCAAG 99
DB 6 AGAAGATGTGCGGAGGAGGAGGAAACAAGCGTTCGCGCGCGCGCGGCGGCGGCGCAAG 65
QY 100 CGCGGTGCGCGCGGAGCGACGCGGCGCTCCGAGAGCGCGCGCGCGCGATATCGTCTGCC 159
DB 66 CGCGAGCGCGCGGAGTACGCGCGCTCCGAAACCGCGCGGAGCGGTACGCTGTAGCC 125
QY 160 CAGATATCGAAGAACAGGAGGTTGCGGTCGCGACCTGGAAACGCGAAGTCTGCTCGAC 219
DB 126 GAGATATCGAAGAACAGGAGGTTGCGTTAGAGCTTGGAAAGCGGCTTACGTCTGAC 185
QY 220 ATCCGCGAGTCTTACGAGAGGAGCGCAAGACCTCCCGCGCGCGCAAGGTATACAGCTC 279
DB 186 CTCGCGAGTCTTACGTCGAGGACGCGCAAGACTCTCCCGCGCGCAAGGTATATCACTC 245
QY 280 CCAATGGATCAGTGGAGATACCTGAGGAGCAATATCAAGCTATAGTAGGGCCATCAAG 339
DB 246 CAGTTAGATCAGTGGAGATATTTGAGGAGCAATATCAAGCCATAGTAGGGCCATCAAG 305
QY 340 GAGAATCGGTATCGGAGCC 359
DB 306 GAGAACACATGATGGAGCC 325
Search completed: January 11, 2005, 22:08:08
Job time : 2763 secs
```

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2005, 22:49:11 ; Search time 152 Seconds  
(without alignments)  
238.366 Million cell updates/sec

Title: US-10-629-953-4  
Perfect score: 530  
Sequence: 1 MWRKGNKRFGGGEPAAKRR.....QWKILRDNKAIDEAIKENA 101

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_23Sep04:.\*  
1: Geneseqp1980s:.\*  
2: Geneseqp1990s:.\*  
3: Geneseqp2000s:.\*  
4: Geneseqp2001s:.\*  
5: Geneseqp2002s:.\*  
6: Geneseqp2003as:.\*  
7: Geneseqp2003bs:.\*  
8: Geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	530	100.0	101	3	AAY44879 Rice tran
2	398	75.1	103	3	AAY44878 Corn tran
3	389.5	73.5	96	3	AAY44881 Wheat tra
4	269	50.8	105	3	AAY44880 Soybean t
5	247.5	46.7	107	3	AAG18457 Arabidops
6	247.5	46.7	142	3	AAG18455 Arabidops
7	206.5	39.0	111	3	AAY44883 Vernonia
8	194.5	36.7	85	3	AAG04501 Arabidops
9	194.5	36.7	120	3	AAG04499 Arabidops
10	185	34.9	165	6	ABR57409 Brassica
11	184.5	34.8	141	3	AAY44886 Soybean t
12	183	34.5	144	3	AAG50454 Arabidops
13	183	34.5	144	3	AAG09966 Arabidops
14	183	34.5	165	3	AAG50453 Arabidops
15	183	34.5	165	3	AAG09965 Arabidops
16	183	34.5	165	5	ABB92888 Herbicida
17	170.5	32.2	99	3	AAY44882 Marigold
18	165	31.1	140	3	AAY44885 Corn tran
19	165	31.1	183	3	AAY44884 Corn tran
20	160.5	30.3	127	4	ABB50267 Polymar
21	160.5	30.3	127	6	ABR47556 Breast ca
22	160.5	30.3	127	7	ADF76608 Novel hum
23	160.5	30.3	127	8	ADL91494 Human imm
24	160.5	30.3	127	8	ADN04676 Anticpori
25	160.5	30.3	132	3	AAB54327 Human pan

26	154.5	29.2	150	4	ABG27269 Novel hum
27	145.5	27.5	110	4	ABB63838 Drosophi1
28	144	27.2	127	8	ADN99783 Novel hum
29	131.5	24.8	128	5	ABP64800 Human pro
30	116	21.9	113	4	ABG23001 Novel hum
31	116	21.9	150	4	ABG26860 Novel hum
32	112.5	21.2	54	8	ADP29939 Human sec
33	112.5	21.2	54	8	ADP30051 Human sec
34	110	20.8	104	5	ABP33375 Human RNA
35	106	20.0	193	4	ABG23003 Novel hum
36	105.5	19.9	208	4	ABG26861 Novel hum
37	100	18.9	234	3	AAG06282 Arabidops
38	100	18.9	245	3	AAG06281 Arabidops
39	98	18.5	72	3	AAG14414 Arabidops
40	98	18.5	95	3	AAG14413 Arabidops
41	82.5	15.6	11877	8	ADL91934 Streptomy
42	82.5	15.6	12199	3	AAY77180 S. venezu
43	79.5	15.0	161	3	AAG06283 Arabidops
44	79.5	15.0	336	3	AAG11256 Arabidops
45	79.5	15.0	339	4	ABG26859 Novel hum

## ALIGNMENTS

RESULT 1  
AAY44879  
ID AAY44879 standard; protein; 101 AA.  
XX  
AC AAY44879;  
XX  
DT 18-MAY-2000 (first entry)  
XX  
DE Rice transcription coactivator PC4(P15) type 1.  
XX  
KW Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 1;  
KW rice; signal mediator; activator; Gal4/VPI6; transgenic plant;  
KW general transcription factor; GTF; transcription initiation complex;  
KW immunological screening; detection; marker; clone rrl.pk0003.a12.  
XX  
OS Oryza sativa.  
XX  
FN WO200005377-A2.  
XX  
PD 03-FEB-2000.  
XX  
PF 21-JUL-1999; 99WO-US016479.  
XX  
PR 22-JUL-1998; 98US-0093687P.  
XX  
(DUPO ) DU PONT DE NEMOURS & CO E I.  
XX  
PI Cahoon RE, Caiami PG, Odell JT, Sakai H, Zhu Q;  
XX  
WP1; 2000-182701/16.  
DR N-PSDB; AAZ50412.  
XX  
PT Novel PC4 transcriptional coactivator polynucleotides and polypeptides  
PT used to alter the level of PC4(P15) type 1 and PC4(P15) type 2  
PT polynucleotides and polypeptides.  
XX  
PS Claim 12; Page 33; 43pp; English.  
XX  
CC The present amino acid sequence is the rice Positive Cofactor 4 (PC4)  
CC transcription coactivator, designated as PC4(P15) type 1. It is isolated  
CC from clone rrl.pk0003.a12, obtained from rrl CDNA library prepared using  
CC rice root two week old developing seedlings. This sequence has 45%  
CC homology to PC4(P15) type 1 from Arabidopsis thaliana (gi 2997684). PC4  
CC molecules functions as a signal mediator between activators like  
CC Gal4/VPI6 and general transcription factors (GTFs) in a transcription  
CC initiation complex. The expression levels of PC4 can be manipulated and  
CC the functional properties of specific transcriptional activators can be  
CC modulated. The PC4 DNA can be used to create transgenic plants with







PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140655P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 28-JUN-1999; 99US-0140951P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 08-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144408P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144844P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 03-AUG-1999; 99US-0146389P.  
 PR 04-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.

PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161922P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 46.7%; Score 247.5; DB 3; Length 107;

Best Local Similarity 56.5%; Pred. No. 2.6e-22;  
 Matches 48; Conservative 12; Mismatches 22; Indels 3; Gaps 1;

QY 15 PAAKRAAGDGESEADDDIVVAQISKNNRRVAVRTWNGKVVVDIREFYKDGKTLPGRK 74

Db 24 PAKKAKPADDD---SDQSDDIVVCNISKNRRVSVRNWNGKIWDIREFYVVDGKTLPGKK 80

Oy 75 GQLPMDQWKILRNIKAIDRAIKE 99  
||| :||| :||| :  
Db 81 GISLSDQWNLTRNHAEDIERALSD 105

RESULT 6  
AAG18455  
ID AAG18455 standard; protein; 142 AA.  
XX  
AC AAG18455;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 19871.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 02-JUL-1999; 99US-0142154P.  
PR 06-JUL-1999; 99US-0142055P.  
PR 08-JUL-1999; 99US-0142390P.  
PR 09-JUL-1999; 99US-0142803P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145513P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147018P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147835P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.

PR	13-AUG-1999;	99US-0148684P.	
PR	16-AUG-1999;	99US-0149368P.	
PR	17-AUG-1999;	99US-0149175P.	
PR	18-AUG-1999;	99US-0149426P.	
PR	20-AUG-1999;	99US-0149722P.	
PR	20-AUG-1999;	99US-0149723P.	
PR	20-AUG-1999;	99US-0149929P.	
PR	23-AUG-1999;	99US-0149902P.	
PR	23-AUG-1999;	99US-0149930P.	
PR	23-AUG-1999;	99US-0150566P.	
PR	26-AUG-1999;	99US-0150884P.	
PR	27-AUG-1999;	99US-0151065P.	
PR	27-AUG-1999;	99US-0151066P.	
PR	27-AUG-1999;	99US-0151080P.	
PR	30-AUG-1999;	99US-0151303P.	
PR	31-AUG-1999;	99US-0151438P.	
PR	01-SEP-1999;	99US-0151930P.	
PR	07-SEP-1999;	99US-0152363P.	
PR	10-SEP-1999;	99US-0153070P.	
PR	13-SEP-1999;	99US-0153758P.	
PR	15-SEP-1999;	99US-0154018P.	
PR	16-SEP-1999;	99US-0154039P.	
PR	20-SEP-1999;	99US-0154779P.	
PR	22-SEP-1999;	99US-0155139P.	
PR	23-SEP-1999;	99US-0155486P.	
PR	24-SEP-1999;	99US-0155659P.	
PR	28-SEP-1999;	99US-0156458P.	
PR	29-SEP-1999;	99US-0156596P.	
PR	04-OCT-1999;	99US-0157117P.	
PR	05-OCT-1999;	99US-0157753P.	
PR	06-OCT-1999;	99US-0157865P.	
PR	07-OCT-1999;	99US-0158029P.	
PR	08-OCT-1999;	99US-0158322P.	
PR	12-OCT-1999;	99US-0158369P.	
PR	13-OCT-1999;	99US-0159293P.	
PR	13-OCT-1999;	99US-0159294P.	
PR	13-OCT-1999;	99US-0159295P.	
PR	14-OCT-1999;	99US-0159329P.	
PR	14-OCT-1999;	99US-0159330P.	
PR	14-OCT-1999;	99US-0159331P.	
PR	14-OCT-1999;	99US-0159637P.	
PR	14-OCT-1999;	99US-0159638P.	
PR	18-OCT-1999;	99US-0159584P.	
PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161992P.	
PR	28-OCT-1999;	99US-0161993P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match 46.7%; Score 247.5; DB 3; Length 142;			
Best Local Similarity 56.5%; Pred. No. 3.8e-22;			
Matches 48; Conservative 12; Mismatches 22; Indels 3; Gaps 1;			
Qy	15	PAKRAAGDGGSEADDDIWAQISKRRVAVRTWNGKVVDIREFYKDGKTLPGRK	74
Db	59	PAKRAAGDGGSEADDDIWAQISKRRVAVRTWNGKVVDIREFYKDGKTLPGRK	74
Qy	75	GIQLPMDQWKILRDNKAIDEAIKE	99

Db	116	GISLSVDQNTLNHAEDIEKALSD	140
RESULT 7			
AA44883	ID	AA44883	standard; protein; 111 AA.
XX	AC	AA44883;	
XX	DT	18-MAY-2000	(first entry)
XX	DE	Vernonia	transcription coactivator PC4(P15) type 1.
XX	KW	Positive cofactor 4; PC4;	transcription coactivator; PC4(P15) type 1;
XX	KW	vernonia; signal mediator; activator;	Ga14/Vp16; transgenic plant;
XX	KW	general transcription factor; GTF;	transcription initiation complex;
XX	KW	immunological screening; detection;	marker; clone vsln.pk013.f21.
OS	XX	Vernonia	mespilifolia.
XX	XX	Key	Location/Qualifiers
FT	Misc-difference 63	/label= Unknown	
FT	Misc-difference 74	/note= "Encoded by AAN"	
FT	Misc-difference 74	/label= Unknown	
FT	Misc-difference 79	/note= "Encoded by GNG"	
FT	Misc-difference 79	/label= Unknown	
FT	Misc-difference 81	/note= "Encoded by NAC"	
FT	Misc-difference 81	/label= Unknown	
FT	Misc-difference 82	/note= "Encoded by AAN"	
FT	Misc-difference 85	/label= Unknown	
FT	Misc-difference 85	/note= "Encoded by NAA"	
FT	Misc-difference 87	/label= Unknown	
FT	Misc-difference 87	/note= "Encoded by GNC"	
FT	Misc-difference 95	/label= Unknown	
FT	Misc-difference 95	/note= "Encoded by ANA"	
FT	Misc-difference 106	/label= Unknown	
FT	Misc-difference 106	/note= "Encoded by CAN"	
FT	Misc-difference 106	/label= Unknown	
FT	Misc-difference 106	/note= "Encoded by ANT"	
XX	XX	WO200005377-A2.	
XX	XX	03-FEB-2000.	
XX	XX	21-JUL-1999;	99WO-US016479.
XX	XX	22-JUL-1999;	98US-0093687P.
XX	XX	(DUPO ) DU PONT DE NEMOURS & CO E I.	
XX	XX	Cahoon RE, Caimi PG, Odell JT, Sakai H, Zhu Q;	
XX	XX	WPI; 2000-182701/16..	
XX	XX	N-PSDB; AA250416.	
XX	XX	Novel PC4 transcriptional coactivator polynucleotides and polypeptides	
XX	XX	used to alter the level of PC4(P15) type 1 and PC4(P15) type 2	
XX	XX	polynucleotides and polypeptides.	
XX	XX	Claim 6; Page 39; 43pp; English.	
XX	XX	The present amino acid sequence is the vernonia Positive Cofactor 4 (PC4)	
XX	XX	transcription coactivator, designated as PC4(P15) type 1. It is isolated	
XX	XX	from clone vsln.pk013.f21, obtained from vsln cDNA library, prepared from	

```

CC vernonia seeds. This sequence has 41% homology to PC4(P15) type 1 from
CC Arabidopsis thaliana (gi 2997684). PC4 molecules functions as a signal
CC mediator between activators like Gal4/VP16 and general transcription
CC factors (GTFs) in a transcription initiation complex. The expression
CC levels of PC4 can be manipulated and the functional properties of
CC specific transcriptional activators can be modulated. The PC4 DNA can be
CC used to create transgenic plants with altered PC4 levels, that would
CC affect the level of transcription of specific genes in the plant. It is
CC also used for immunological screening of cDNA libraries and to raise
CC specific antibodies for detection. The DNA is used as probes and primers,
CC for genetic and physical mapping of genes and as markers for traits
CC linked to those genes
XX
XX SQ Sequence 111 AA;
XX
XX Query Match 39.0%; Score 206.5; DB 3; Length 111;
XX Best Local Similarity 49.4%; Pred. No. 3.2e-17;
XX Matches 40; Conservative 12; Mismatches 26; Indels 3; Gaps 1;
XX
Oy 17 AKRRAAGDGPESADDDIVVAQISKRRVAVRTWNGKVVDIREFYKDGKTLPGRKGI 76
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 33 SKDTAGDSDE--DDIFICDVSKNRRVSRXWQGRVFDIRXFMKXGXMPXXKGI 89
Oy 77 QLPMDQWKILRDNIKATDEAI 97
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 90 SLTWDXWKLRAHVDEXDKAL 110

RESULT 8
AAG04501
ID AAG04501 standard; protein; 85 AA.
XX
AC AAG04501;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 551.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 18-JUN-1999; 99US-0139763P.
XX 21-JUN-1999; 99US-0139817P.
XX 22-JUN-1999; 99US-0139899P.
XX 23-JUN-1999; 99US-0140353P.
XX 23-JUN-1999; 99US-0140354P.
XX 24-JUN-1999; 99US-0140695P.
XX 28-JUN-1999; 99US-0140823P.
XX 29-JUN-1999; 99US-0140991P.
XX 30-JUN-1999; 99US-0141287P.
XX 01-JUL-1999; 99US-0141842P.
XX 02-JUL-1999; 99US-0142055P.
XX 06-JUL-1999; 99US-0142390P.
XX 08-JUL-1999; 99US-0142803P.
XX 09-JUL-1999; 99US-0142920P.
XX 12-JUL-1999; 99US-0142977P.
XX 13-JUL-1999; 99US-0143542P.
XX 14-JUL-1999; 99US-0143624P.
XX 15-JUL-1999; 99US-0144005P.
XX 16-JUL-1999; 99US-0144085P.
XX 16-JUL-1999; 99US-0144086P.
XX 19-JUL-1999; 99US-0144325P.
XX 19-JUL-1999; 99US-0144331P.
XX 19-JUL-1999; 99US-0144332P.
XX 19-JUL-1999; 99US-0144333P.
XX 19-JUL-1999; 99US-0144334P.
XX 19-JUL-1999; 99US-0144335P.
XX 20-JUL-1999; 99US-0144352P.
XX 20-JUL-1999; 99US-0144632P.
XX 20-JUL-1999; 99US-0144884P.
XX 21-JUL-1999; 99US-0144814P.
XX 21-JUL-1999; 99US-0145086P.
XX 21-JUL-1999; 99US-0145088P.
XX 22-JUL-1999; 99US-0145085P.
XX 22-JUL-1999; 99US-0145087P.
XX 22-JUL-1999; 99US-0145089P.
XX 22-JUL-1999; 99US-0145192P.
XX 23-JUL-1999; 99US-0145145P.
XX 23-JUL-1999; 99US-0145218P.

```

PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147152P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147433P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 16-AUG-1999; 99US-0148684P.  
 PR 17-AUG-1999; 99US-0149368P.  
 PR 18-AUG-1999; 99US-0149175P.  
 PR 20-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149743P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150556P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151085P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 16-SEP-1999; 99US-0154018P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.

PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161922P.  
 PR 28-OCT-1999; 99US-0161933P.  
 PR 23-OCT-1999; 99US-0162142P.

Query Match 36.7%; Score 194.5; DB 3; Length 85;  
 -Best Local Similarity 63.9%; Pred. No. 6.8e-16;  
 Matches 39; Conservative 4; Mismatches 15; Indels 3; Gaps 1;

QY 15 PAAKRAAGDDGSESDDDIVVAQISKRRVAVRTWNGKVVVDIREFYKDGKTLPGRK 74  
 DB 24 PAKKVPKPADDD---SDQSDDIVVCNISKNRRSVRNWNGKIWDIREFYKDGKTLPGKK 80

QY 75 G 75  
 DB 81 G 81

RESULT 9  
 AAG04499  
 ID AAG04499 standard; protein; 120 AA.  
 AC AAG04499;  
 XX  
 DT 17-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 549.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PE 25-FEB-2000; 2000EP-00301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 03-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 23-APR-1999; 99US-0130891P.  
 PR 28-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 30-APR-1999; 99US-0132407P.  
 PR 04-MAY-1999; 99US-0132484P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 06-MAY-1999; 99US-0132487P.  
 PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0134256P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.



PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161922P.  
 PR 28-OCT-1999; 99US-0161933P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 36.7%; Score 194.5; DB 3; Length 120;  
 Best Local Similarity 63.9%; Pred. No. 1.1e-15;  
 Matches 39; Conservative 4; Mismatches 15; Indels 3; Gaps 1;

QY 15 PAAKRAAGDDGPGSEADDDIVVAQISKNRVAVRTWNGKVVDIREFYEKDGKTLPGRK 74  
 Db 59 PAKVAKPADDD---SQSDDIVVCNISKNRVSVRNWNGKINIDIREFVYKDGKTLPGKK 115

QY 75 G 75  
 Db 116 G 116

RESULT 10  
 ABR57409  
 ID ABR57409 standard; protein; 165 AA.  
 AC ABR57409;  
 XX  
 XX 23-OCT-2003 (revised)  
 DT 12-SEP-2003 (first entry)  
 XX  
 XX Brassica campestris MIP102 protein SEQ ID NO:2.  
 XX  
 XX Brassica campestris; plant; MIP102; virus resistance;  
 KW plant virus transport protein binding protein.  
 XX  
 XX Brassica rapa.  
 OS  
 XX WO2003022039-A1.  
 PN  
 PD 20-MAR-2003.  
 XX  
 XX 10-SEP-2001; 2001WO-JP007858.  
 PF  
 XX 10-SEP-2001; 2001WO-JP007858.  
 PR  
 XX (NAAG-) NAT INST AGROBIOLOGICAL SCI.  
 PA  
 XX Nishiguchi M, Nyunoya H, Matsushita Y;  
 PI WPI; 2003-313169/30.  
 DR N-PSDB; ACF03526.  
 DR  
 XX Plant protein binding to plant virus transport protein for imparting  
 PT virus resistance to plants including tobacco.  
 XX  
 XX Claim 2; Fig 2; 84pp; Japanese.  
 PS  
 XX The present invention describes a method for imparting virus resistance  
 CC to plants in which the plant cells are transformed with a polynucleotide  
 CC encoding a protein (I) which binds to plant virus transport protein. Also  
 CC described are plants transformed by (I). The method is useful for  
 CC increasing resistance of plants such as tobacco, tomato, petunia, pear,  
 CC rice, wheat, barley, maize, soybean, oilseed rape, rose, apple, alfalfa,  
 CC melon, lavender, onion, spinach and parsley; to viruses including  
 CC tobamovirus, tobamovirus, dianthovirus, alfamovirus, bromovirus,  
 CC cucumovirus, comovirus, nepovirus, caulimovirus, geminivirus, potyvirus  
 CC and tospovirus. The present sequence represents Brassica campestris  
 CC MIP102, which is a specifically claimed plant virus transport protein  
 CC binding protein from the present invention.. (Updated on 23-OCT-2003 to  
 CC standardise OS field)  
 XX  
 XX Sequence 165 AA;

Query Match 34.9%; Score 185; DB 6; Length 165;  
 Best Local Similarity 38.1%; Pred. No. 2.5e-14;  
 Matches 37; Conservative 18; Mismatches 22; Indels 20; Gaps 2;

QY 3 RKGNKRFGGEPAAKRAAGDDGPGSEADDDIVVAQISKNRVAVRTWNGKVVDIREF 62  
 Db 84 KEGNKEF-----DD-----DGLIICRLSDKRRVTIQEPRGKSLVSIREY 123

QY 63 YEKDGKTLPGKGIQLPMDQWKLIRDNIIKAIDEAIKE 99  
 Db 124 YKDGKELPSKSGISLTDQWSTFKKNIPAI EAIAVKK 160

RESULT 11  
 AAY44886  
 ID AAY44886 standard; protein; 141 AA.  
 XX  
 XX AAY44886;  
 AC  
 XX 18-MAY-2000 (first entry)  
 DT  
 XX Soybean transcription coactivator PC4 (P15) type 2.  
 DE  
 XX Positive cofactor 4; PC4; transcription coactivator; PC4 (P15) type 2;  
 KW soybean; signal mediator; activator; Gal4/VP16; transgenic plant;  
 KW general transcription factor; GFP; transcription initiation complex;  
 KW immunological screening; detection; marker; clone ses4d.pk0016.g2.  
 XX  
 XX Glycine max.  
 OS  
 XX WO200005377-A2.  
 PN  
 XX 03-FEB-2000.  
 PD  
 XX 21-JUL-1999; 99WO-US016479.  
 PF  
 XX 22-JUL-1998; 98US-0093687P.  
 PR  
 XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
 PA  
 XX Cahoon RE, Caimi PG, Odell JT, Sakai H, Zhu Q;  
 PI WPI; 2000-182701/16.  
 DR N-PSDB; AA250419.  
 DR  
 XX Novel PC4 transcriptional coactivator polynucleotides and polypeptides  
 PT used to alter the level of PC4 (P15) type 1 and PC4 (P15) type 2  
 PT polynucleotides and polypeptides.  
 XX  
 XX Claim 18; Page 42; 43pp; English.  
 PS  
 XX The present amino acid sequence is the soybean Positive Cofactor 4 (PC4)  
 CC transcription coactivator, designated as PC4 (P15) type 2. It is isolated  
 CC from clone ses4d.pk0016.g2, obtained from ses4d cDNA library, prepared  
 CC from soybean embryogenic suspension 4 days after subculture. This  
 CC sequence has 66% homology to PC4 (P15) type 2 from Arabidopsis thaliana  
 CC (gi 2997686). PC4 molecules functions as a signal mediator between  
 CC activators like Gal4/VP16 and general transcription factors (GTFs) in a  
 CC transcription initiation complex. The expression levels of PC4 can be  
 CC manipulated and the functional properties of specific transcriptional  
 CC activators can be modulated. The PC4 DNA can be used to create transgenic  
 CC plants with altered PC4 levels, that would affect the level of  
 CC transcription of specific genes in the plant. It is also used for  
 CC immunological screening of cDNA libraries and to raise specific  
 CC antibodies for detection. The DNA is used as probes and primers, for  
 CC genetic and physical mapping of genes and as markers for traits linked to  
 CC those genes  
 XX  
 XX Sequence 141 AA;

Query Match 34.8%; Score 184.5; DB 3; Length 141;  
 Best Local Similarity 38.7%; Pred. No. 2.3e-14;



Matches	36; Conservative	21; Mismatches	25; Indels	11; Gaps	2;
Qy	7	KRFGGGEPAARRAAGDSEBSADDDIVVAQISKNRRVAVRTWNGKVVVDIREFYKDD	66		
Db	56	EELGGG---SKKEYDEG-----DLIICRLSDKRRVTIQDFRGKTLVSIREYKDD	104		
Qy	67	GKTLPGRGIGIOLPMDONKILRDNIKAIDEATKE	99		
Db	105	GKELPTSGKISLTTEQMSAFKKNVPAYEKAIKK	137		
RESULT 12					
AAG50454					
ID	AAG50454 standard; protein; 144 AA.				
XX	AAG50454;				
AC					
XX	18-OCT-2000 (first entry)				
DT					
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 63942.				
DE					
XX	Protein identification; signal transduction pathway; metabolic pathway;				
KW	hybridisation assay; genetic mapping; gene expression control; promoter;				
KW	termination sequence.				
XX	Arabidopsis thaliana.				
OS					
XX	EP1033405-A2.				
PN					
XX	06-SEP-2000.				
PD					
XX					
PF	25-FEB-2000; 2000EP-00301439.				
XX					
PR	25-FEB-1999;	99US-0121825P.	PR	10-JUN-1999;	99US-0138540P.
PR	05-MAR-1999;	99US-0123180P.	PR	10-JUN-1999;	99US-0138847P.
PR	09-MAR-1999;	99US-0123548P.	PR	14-JUN-1999;	99US-0139119P.
PR	23-MAR-1999;	99US-0125788P.	PR	16-JUN-1999;	99US-0139452P.
PR	25-MAR-1999;	99US-0126264P.	PR	16-JUN-1999;	99US-0139453P.
PR	29-MAR-1999;	99US-0126785P.	PR	17-JUN-1999;	99US-0139492P.
PR	01-APR-1999;	99US-0127462P.	PR	18-JUN-1999;	99US-0139454P.
PR	06-APR-1999;	99US-0128234P.	PR	18-JUN-1999;	99US-0139455P.
PR	08-APR-1999;	99US-0128714P.	PR	18-JUN-1999;	99US-0139456P.
PR	16-APR-1999;	99US-0129845P.	PR	18-JUN-1999;	99US-0139457P.
PR	19-APR-1999;	99US-0130077P.	PR	18-JUN-1999;	99US-0139458P.
PR	21-APR-1999;	99US-0130449P.	PR	18-JUN-1999;	99US-0139459P.
PR	23-APR-1999;	99US-0130510P.	PR	18-JUN-1999;	99US-0139460P.
PR	23-APR-1999;	99US-0130891P.	PR	18-JUN-1999;	99US-0139461P.
PR	28-APR-1999;	99US-0131449P.	PR	18-JUN-1999;	99US-0139462P.
PR	30-APR-1999;	99US-0132048P.	PR	18-JUN-1999;	99US-0139463P.
PR	30-APR-1999;	99US-0132407P.	PR	18-JUN-1999;	99US-0139464P.
PR	04-MAY-1999;	99US-0132484P.	PR	18-JUN-1999;	99US-0139465P.
PR	05-MAY-1999;	99US-0132485P.	PR	18-JUN-1999;	99US-0139466P.
PR	06-MAY-1999;	99US-0132486P.	PR	18-JUN-1999;	99US-0139467P.
PR	06-MAY-1999;	99US-0132487P.	PR	18-JUN-1999;	99US-0139468P.
PR	07-MAY-1999;	99US-0132863P.	PR	18-JUN-1999;	99US-0139469P.
PR	11-MAY-1999;	99US-0133256P.	PR	18-JUN-1999;	99US-0139470P.
PR	14-MAY-1999;	99US-0134218P.	PR	18-JUN-1999;	99US-0139471P.
PR	14-MAY-1999;	99US-0134219P.	PR	18-JUN-1999;	99US-0139472P.
PR	14-MAY-1999;	99US-0134221P.	PR	18-JUN-1999;	99US-0139473P.
PR	14-MAY-1999;	99US-0134370P.	PR	18-JUN-1999;	99US-0139474P.
PR	18-MAY-1999;	99US-0134768P.	PR	18-JUN-1999;	99US-0139475P.
PR	19-MAY-1999;	99US-0134941P.	PR	18-JUN-1999;	99US-0139476P.
PR	20-MAY-1999;	99US-0135124P.	PR	18-JUN-1999;	99US-0139477P.
PR	21-MAY-1999;	99US-0135353P.	PR	18-JUN-1999;	99US-0139478P.
PR	24-MAY-1999;	99US-0135629P.	PR	18-JUN-1999;	99US-0139479P.
PR	25-MAY-1999;	99US-0136021P.	PR	18-JUN-1999;	99US-0139480P.
PR	27-MAY-1999;	99US-0136392P.	PR	18-JUN-1999;	99US-0139481P.
PR	28-MAY-1999;	99US-0136782P.	PR	18-JUN-1999;	99US-0139482P.
PR	01-JUN-1999;	99US-0137222P.	PR	18-JUN-1999;	99US-0139483P.
PR	03-JUN-1999;	99US-0137528P.	PR	18-JUN-1999;	99US-0139484P.
PR	04-JUN-1999;	99US-0137502P.	PR	18-JUN-1999;	99US-0139485P.
PR	07-JUN-1999;	99US-0137724P.	PR	18-JUN-1999;	99US-0139486P.
PR	08-JUN-1999;	99US-0138094P.	PR	18-JUN-1999;	99US-0139487P.



PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139452P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.

PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 20-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151067P.  
 PR 30-AUG-1999; 99US-0151130P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 22-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 28-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 34.5%; Score 183; DB 3; Length 144;

Best Local Similarity 38.5%; Pred. No. 3,7e-14;

Matches 37; Conservative 17; Mismatches 22; Indels 20; Gaps 2;

Qy 4 KGNKFCGGGEPKRRACGGDGPSESADDDIVVAQISKRRVAVRTWNGKVVVDIREFY 63

Db 64 KGNKEF-----DD-----DGLIFCLSKRRVTIQEFKGLVSIREFY 103



```
PR 16-AUG-1999; 99US-0149358P.
PR 17-AUG-1999; 99US-01491175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-015753P.
PR 06-OCT-1999; 99US-0157855P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159284P.
PR 13-OCT-1999; 99US-0159255P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 34.5%; Score 183; DB 3; Length 165;
Best Local Similarity 38.5%; Pred. No. 4.4e-14;
Matches 37; Conservative 17; Mismatches 22; Indels 20; Gaps 2;

Qy 4 KGNKFGGGEPAKRAAGDGPESADDDIVVAQISKNRVAVRTWNGKVVVDIREFY 63
Db 85 KGNKEF-----DD-----DGLIICRLSDKERVITQEFKGLSVIREY 124

Qy 64 EKQKTLPGKRGIGIOLPMDQWKLKLDNIKAIDEATKE 99
Db 125 KDGKGLPSTKGLISLTDEQWSTFKKNMPATENAVK 160
```

RESULT 15  
AAG09965  
ID AAG09965 standard; protein; 165 AA.  
XX  
AC AAG09965;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 8101.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; Genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125768P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 06-APR-1999; 99US-0127462P.  
PR 08-APR-1999; 99US-0128234P.  
PR 16-APR-1999; 99US-0128714P.  
PR 19-APR-1999; 99US-0129845P.  
PR 21-APR-1999; 99US-0130077P.  
PR 23-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139452P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.

PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 24-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-014091P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142300P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 18-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145227P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.

PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 34.5%; Score 183; DB 3; Length 165;  
Best Local Similarity 38.5%; Pred. No. 4.4e-14;  
Matches 37; Conservative 17; Mismatches 22; Indels 20; Gaps 2;

QY 4 KGNKRGGGGPAKRAAGDDGPFSESADDDIVVAGISKNRRYAVRTWNGKVVVDIREFY 63  
DB 85 KGNKEF-----DD-----DGDLLICLSDKRRVTIOEFKGSLSIREYY 124

QY 64 EKDGKTLPGKKGQLPMDOWKILRDNKKAIDEAIKE 99

DB 125 KDGKELPTSIGISLTDQWSTFKNMPAENAVKK 160

Search completed: January 10, 2005, 23:03:29

Job time : 156 secs

-----

2005-01-13 09:21:49

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**



A;Residues: 1-165 <COR>  
A:Cross-references: UNIPROT:065155: EMBL:AF053303: PIDN:AAC08575.1



Df 65 YYLDSSTGEMKPGKGISLTEDLYDLLKHLRLNIDEALR 103

RESULT 8  
S66164  
single strand DNA-binding protein C31A - fruit fly (*Drosophila melanogaster*)  
C;Species: *Drosophila melanogaster*  
C;Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S66164  
R;Lacoste, J.; Codani-Simonart, S.; Best-Belpomme, M.; Peronnet, F.  
Nucleic Acids Res. 23, 5073-5079, 1995  
A;Title: Characterization and cloning of p1l, a transrepressor of *Drosophila melanogaster*  
A;Reference number: S66164; MUID:96145266; PMID:8559667  
A;Accession: S66164  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-98 <AC>  
A;Cross-references: UNIPROT:Q24026; EMBL:U18971; NID:g1109661; PIDN:AAC46989.1; PID:g1109661

Query Match 21.1%; Score 112; DB 2; Length 98;  
Best Local Similarity 35.7%; Pred. No. 0.00049;  
Matches 25; Conservative 15; Mismatches 26; Indels 4; Gaps 2;

Qy 14 EPAKRRRAAGD---DGSSADDIIVQAISKNRVAVRTWNGKVVDIREFYEKDKGKTL 70  
Db 22 KPASKAKESDPNSDPKDSGENGATSWTLEGILRVINEFRGRKSVDIRFYEYDKGGQIL 81  
Qy 71 PGRKGIQLPM 80  
Db 82 PGKKG-HLPI 90

RESULT 9  
T24876  
hypothetical protein T13F2.2 - *Caenorhabditis elegans*  
C;Species: *Caenorhabditis elegans*  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T24876  
R;Swinnburne, J.  
submitted to the EMBL Data Library, October 1996  
A;Reference number: Z19947  
A;Accession: T24876  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-124 <WL>  
A;Cross-references: UNIPROT:Q94045; EMBL:Z81122; PIDN:CAB03353.1; GSPDB:GN00022; CRSP:T13F2  
C;Genetics:  
A;Gene: CRSP:T13F2.2  
A;Map position: 4  
A;Introns: 27/3; 61/3; 91/3  
C;Superfamily: *Caenorhabditis elegans* hypothetical protein T13F2.2

Query Match 19.2%; Score 102; DB 2; Length 124;  
Best Local Similarity 42.4%; Pred. No. 0.0061;  
Matches 25; Conservative 9; Mismatches 23; Indels 2; Gaps 2;

Qy 39 QISKNRRAVRTWNGKVVDIREFY-EKDQ-KYLPGRKGIQLPMDQWKILRDNIKAIDE 95  
Db 64 EIGNLRATVTSKFGKEYYNIREYYIDRSQKMPSRGKISLSKAQWANLKOLIPSIDK 122

RESULT 10  
T47393  
hypothetical protein T18D12.80 - *Arabidopsis thaliana*  
C;Species: *Arabidopsis thaliana* (mouse-ear cress)  
C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
C;Accession: T47393  
R;Massenet, O.; Clabault, G.; Quigley, F.; Maché, R.; Mewes, H.W.; Lemcke, K.; Mayer, K.  
submitted to the Protein Sequence Database, March 2000  
A;Reference number: Z24464  
A;Accession: T47393

A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-421 <MAS>  
A;Cross-references: UNIPROT:Q9W239; EMBL:ALJ138644  
A;Experimental source: cultivar Columbia; BAC clone T18D12  
C;Genetics:  
A;Map position: 3  
A;Introns: 66/3; 101/1; 143/2  
A;Note: T18D12.80

Query Match 16.4%; Score 87; DB 2; Length 421;  
Best Local Similarity 27.3%; Pred. No. 0.69;  
Matches 27; Conservative 16; Mismatches 34; Indels 22; Gaps 3;

QY 15 PAAKRAAGDGPESADDDIVVAQISKNRVAVRTWNGKVVVDIREFY---EKDGG--- 68  
DB 49 PAYKRVKS-----VAEERFLYKSETKDVWVKNYQGTSLIGKGVKCKDAKFA 101  
QY 69 -----TLPGRKGQLPMDQWKILRDNKAIDEAIK 98  
DB 102 TLVNDERKAFMLCSLGLISFPHQSVFKNFSAITEAIK 140

RESULT 11  
D49348  
succinoglycan biosynthesis glycosyltransferase (EC 2.4.1.-) exoU [validated] - Rhizobium  
C;Species: Rhizobium meliloti  
C;Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 09-Jul-2004  
R;Gluckemann, M.A.; Reuber, T.L.; Walker, G.C.  
J. Bacteriol. 175, 7033-7044, 1993  
A;Title: Family of glycosyl transferases needed for the synthesis of succinoglycan by Rh  
A;Reference number: A49348; MUID:94042869; PMID:8226645  
A;Accession: D49348  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-342 <GLU>  
A;Cross-references: UNIPROT:P33700; GB:L20758; NID:g393240; PIDN:AAA16053.1; PID:g393252  
A;Genetics:  
A;Gene: exoU  
C;Function:  
A;Description: EC 2.4.1.-; succinoglycan biosynthesis glycosyltransferase [validated, MU  
exoU, and exoW together are responsible for sugar addition to the lipid carrier  
A;Pathway: succinoglycan biosynthesis  
C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 15.3%; Score 81; DB 2; Length 342;  
Best Local Similarity 27.9%; Pred. No. 2.1;  
Matches 29; Conservative 9; Mismatches 36; Indels 30; Gaps 3;

QY 14 EPAKRRRAAGDGPESADDDIVVAQ-----SADDD---IVVAQISKNR-----VAVRTWNGKV 52  
DB 32 EPEAAEVVVDDG---STDDASVARAADDGTGRLNVVFEENRGPAARHAIAISHSP 88  
QY 53 GKVVVDIREFYKDGKTLPGKGIQLPMDQWKILRDNKAIDEA 96  
DB 89 LIGVLDAADDF-----FPGRLQLLSQDGDWDFIADNFIADAA 126

RESULT 12  
G95975  
glucosyltransferase protein [imported] - Sinorhizobium meliloti (strain 1021) magaplasmi  
C;Species: Sinorhizobium meliloti  
C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A;Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endo  
A;Reference number: A95842; MUID:21396508; PMID:11481431  
A;Accession: G95975  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-342 <RUC>

A;Cross-references: UNIPROT:P33700; GB:AL591985; PIDN:CAC49471.1; PID:g15140957; GSPDB:G  
A;Experimental source: strain 1021, megaplasmid pSymB  
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
L.; Hyman, R.W.; Jones, I.  
Science 293, 668-672, 2001  
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
hebaute, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A;Reference number: A96039; MUID:21368234; PMID:11474104  
A;Contents: annotation  
C;Genetics:  
A;Gene: exoU; SMB20948  
A;Genome: plasmid

Query Match 15.3%; Score 81; DB 2; Length 342;  
Best Local Similarity 27.9%; Pred. No. 2.1;  
Matches 29; Conservative 9; Mismatches 36; Indels 30; Gaps 3;

QY 14 EPAKRRRAAGDGPESADDDIVVAQ-----SADDD---IVVAQISKNR-----VAVRTWNGKV 52  
DB 32 EPEAAEVVVDDG---STDDASVARAADDGTGRLNVVFEENRGPAARHAIAISHSP 88  
QY 53 GKVVVDIREFYKDGKTLPGKGIQLPMDQWKILRDNKAIDEA 96  
DB 89 LIGVLDAADDF-----FPGRLQLLSQDGDWDFIADNFIADAA 126

RESULT 13  
S40173  
ExoU protein - Rhizobium meliloti  
C;Species: Rhizobium meliloti  
C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 06-Jan-1995  
C;Accession: S40173  
R;Becker, A.; Kleickmann, A.; Kuester, H.; Keller, M.; Arnold, W.; Puhler, A.  
submitted to the EMBL Data Library, April 1993  
A;Description: Analysis of the Rhizobium meliloti genes exoU, exoV, exoW, exoT and exoI  
ses.  
A;Reference number: S40173  
A;Accession: S40173  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-288 <BEC>  
A;Cross-references: EMBL:Z22646

Query Match 14.6%; Score 77.5; DB 2; Length 288;  
Best Local Similarity 29.0%; Pred. No. 3.9;  
Matches 29; Conservative 12; Mismatches 36; Indels 23; Gaps 4;

QY 14 EPAKRRRAAGDGPSE-----SADDD---IVVAQISKNR-----VAVRTWNGKV 56  
DB 32 EPEAAEVVVDDGSDTDDASVARAADDGTGRLNVVFEENRGPAARQAISHSPLIGV 91  
QY 57 VDIRFYKDGKTLPGKGIQLPMDQWKILRDNKAIDEA 96  
DB 92 LDADDDFF-----FPGRLQLLSQDGDWDFIADNFIADAA 125

RESULT 14  
A11680  
dihydrofolate reductases homolog lin1987 [imported] - Listeria innocua (strain Clip11262)  
C;Species: Listeria innocua  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 16-Aug-2004  
C;Accession: A11680  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,  
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duseurget, O.; Entian, K.D.; Feihl, H.;  
Science 294, 849-852, 2001  
A;Authors: Krefte, J.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
ok, C.; Schluter, T.; Simoes, N.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mat  
A;Title: Comparative genomics of Listeria species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: A11680

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-160 <GLA>  
A:Cross-references: UNIPROT:Q92AD5; GB:ALS92022; PIDN:CAC97217.1; PID:gi6414488; GSPDB:G  
A:Experimental source: strain Clip11262  
C:Genetics:  
A:Gene: lin1987  
C:Superfamily: Dihydrofolate reductase; type I dihydrofolate reductase homology

Query Match 14.3%; Score 76; DB 2; Length 160;  
Best Local Similarity 40.0%; Pred. No. 2.8;  
Matches 22; Conservative 8; Mismatches 17; Indels 8; Gaps 2;  
Oy 49 RTWNGKVVDIREFYEKDGKTLPGKGI-----QLPNDQWKIL--RDNIKAIDE 95  
Db 33 KTTGKTLVMGRKTYESLGKALPNRKTIVLTRDNEQLQDADAEILHSRDEVILALAE 87

RESULT 15  
D84828  
AP2 domain transcription factor [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: D84828  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: D84828  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-177 <STO>  
A:Cross-references: UNIPROT:Q9SIZ0; GB:AE002093; NID:g4586052; PIDN:AAD25670.1; GSPDB:GN  
A:Gene: Atg40350  
A:Map position: 2

Query Match 14.1%; Score 74.5; DB 2; Length 177;  
Best Local Similarity 30.4%; Pred. No. 4.4;  
Matches 34; Conservative 11; Mismatches 32; Indels 35; Gaps 7;  
Oy 2 WRKGNKRF-----GGGGEPAK---RRA-----AGDDGPSESADDDIVVAQISKRRV 46  
Db 20 WREYNEQTEADSCIDGGGSKPIRKAPPKRSKRGCKGKGPENGICDYTGVRQ----- 72  
Oy 47 AVRTWNGKVVDIREFYEKDGKTLPGKGIQLPNDQWKILRDNIKAIDEAIK 98  
Db 73 --RTW-GKVAEIRE-----PGR-GAKLWLGTFSSSYEALAYDEASK 111

Search completed: January 10, 2005, 23:07:31  
Job time : 41 secs

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2005, 22:49:56 ; Search time 193 Seconds  
(without alignments)  
301.103 Million cell updates/sec

Title: US-10-629-953-4

Perfect score: 530

Sequence: 1 MWRKGNKRFGGGEPAAKRR.....QWKILRNIKAIDEAIKENA 101

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	530	100.0	101	2	Q6Z777 Oryza sativ
2	530	100.0	101	2	BAD15868 Oryza sat
3	247.5	46.7	107	1	KIWI ARATH
4	185	34.9	165	2	Q9AVF8 Arabidopsis
5	183	34.5	165	1	KELP ARATH
6	180.5	34.1	184	2	Q94JF5 Arabidopsis
7	172	32.5	126	1	TCP4 MOUSE
8	172	32.5	127	2	BAC33314 mus muscu
9	172	32.5	127	2	BAC33642 mus muscu
10	166	31.3	119	1	TCP4 RAT
11	165	31.1	110	2	Q9FY30 Arabidopsis
12	161.5	30.5	73	2	Q6E433 homo sapien
13	160.5	30.3	126	1	TCP4 HUMAN
14	160.5	30.3	127	2	CAG33183 homo sapi
15	160	30.2	124	2	Q6DGP6 Arabidopsis
16	153	28.9	128	2	Q6NTX9 Arabidopsis
17	153	28.9	128	2	RAH68824 Arabidopsis
18	151	28.5	128	2	Q6DJR3 Arabidopsis
19	149	28.1	136	2	Q6C485 Arabidopsis
20	145.5	27.5	110	1	TCP4 DROME
21	137	25.8	594	2	Q93YB6 Arabidopsis
22	133.5	25.2	136	1	TCP4 SCHPO
23	133.5	25.2	273	2	Q75DD4 Arabidopsis
24	133.5	25.2	273	2	AA50863 Arabidopsis
25	118	22.3	172	1	TCP4 NEUCR
26	117	22.1	135	2	Q6BPT2 Arabidopsis
27	117	22.1	292	1	SUB1 YEAST
28	112	21.1	98	2	Q240Z6 Arabidopsis
29	112	21.1	277	2	Q6FU60 Arabidopsis
30	102	19.2	124	1	TCP4 CAEEL
31	95	17.9	244	2	Q6CIG4 Arabidopsis

32 95 17.9 409 2 Q6DBG5 Arabidopsis  
33 88 16.6 269 1 PFG VIBPA Arabidopsis  
34 87 16.4 421 2 Q9M239 Arabidopsis  
35 86.5 16.3 488 2 Q8GXX5 Arabidopsis  
36 81 15.3 71 2 Q72LS9 Arabidopsis  
37 81 15.3 71 2 Q8CX1 Arabidopsis  
38 81 15.3 71 2 AAS72010 Arabidopsis  
39 81 15.3 342 1 EXOU RHIME Arabidopsis  
40 81 15.3 1018 2 Q6LOV3 Arabidopsis  
41 80.5 15.2 269 1 PFG VIBV Arabidopsis  
42 80 15.1 3985 2 Q7Q293 Arabidopsis  
43 78.5 14.8 269 1 PFG VIBV Arabidopsis  
44 78.5 14.8 1646 2 Q7PRH5 Arabidopsis  
45 76.5 14.4 271 2 Q8HYL3 Arabidopsis

#### ALIGNMENTS

RESULT 1  
Q6Z777 PRELIMINARY; PRT; 101 AA.  
AC Q6Z777  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Transcriptional coactivator p15 (PC4) family protein-like.  
GN Name=P0470G10.10;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP004876; BAD15868.1;  
DR InterPro; IPR003173; PC4.  
DR InterPro; IPR009044; ssDNA\_bind\_regul.  
DR Pfam; PF02229; PC4; 1  
SQ SEQUENCE 101 AA; 11278 MW; 43AA93C0B488C92 CRC64;

Query Match 100.0%; Score 530; DB 2; Length 101;  
Best Local Similarity 100.0%; Pred. No. 2.9e-45;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWRKGNKRFGGGEPAAKRRAGDGFSESADDDIVVAQISKRRVAVRTWNGKVVDIR 60  
|||  
Db 1 MWRKGNKRFGGGEPAAKRRAGDGFSESADDDIVVAQISKRRVAVRTWNGKVVDIR 60  
|||

Qy 61 EFYEKDGKTLPGKGIQLPMDQWKILRDNKAIKIDEAIKENA 101  
|||  
Db 61 EFYEKDGKTLPGKGIQLPMDQWKILRDNKAIKIDEAIKENA 101  
|||

RESULT 2  
BAD15868 PRELIMINARY; PRT; 101 AA.

ID BAD15868  
AC BAD15868  
DT 23-APR-2004 (TrEMBLrel. 27, Created)  
DT 23-APR-2004 (TrEMBLrel. 27, Last sequence update)  
DT 23-APR-2004 (TrEMBLrel. 27, Last annotation update)  
DE Transcriptional coactivator p15 (PC4) family protein-like.  
GN P0470G10.10;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza; Oryza sativa.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Nipponbare;







```

DR EMBL; AL161518; CAB81193.1; -.
DR EMBL; AY085896; AAM63108.1; -.
DR EMBL; BT005847; AAO64782.1; -.
DR PIR; T52114; T52114.
DR HSP; P53999; 1PCF.
DR GeneFarm; 2402; -.
DR InterPro; IPR003173; PC4.
DR InterPro; IPR009044; sBDNA_bind_regul.
DR Pfam; PF02229; PC4; 1
KW Activator; DNA-binding; Nuclear protein; Transcription;
KW Transcription regulation.
SQ SEQUENCE 165 AA; 19177 MW; 6620F1693D40D3D1 CRC64;
Query Match 34.5%; Score 183; DB 1; Length 165;
Best Local Similarity 38.5%; Pred. No. 2.7e-10;
Matches 37; Conservative 17; Mismatches 22; Indels 20; Gaps 2;
Qy 4 KGNKRGGGGPAKRAAGDGPSESADDIVVAQISKNNRVAVRTWNGKVVDIREFY 63
Db 85 KGNKEF-----DD-----DGLIICRLSDKRRVTIQEFKGLSVSIREY 124
Qy 64 EKDGKTLPGKGIQLPMDQWKILRDNIKAIDEAIKE 99
Db 125 KDGKELPTSGISLTDEQWSTPKKNMPAIEAVK 160
RESULT 6
Q94JES PRELIMINARY; PRT; 184 AA.
AC Q94JES;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE P684B02.23 protein (P0692C11.10 protein).
GN Name=P0684B02.23; Synonym=P0692C11.10;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Niimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Maekawa M., Arikawa K., Chiden Y., Hayaishi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikeno M., Itoh S., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagaaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AP003023; BAB44036.1; -.
DR EMBL; AP003381; BAB86514.1; -.
DR HSP; P53999; 1PCF.
DR Gramene; Q94JES; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003713; F:transcription coactivator activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003173; PC4.
DR InterPro; IPR009044; sBDNA_bind_regul.
DR Pfam; PF02229; PC4; 1.
SQ SEQUENCE 184 AA; 21028 MW; BB1EFAD4C6340B5A CRC64;
Query Match 34.1%; Score 180.5; DB 2; Length 184;
Best Local Similarity 36.5%; Pred. No. 5.4e-10;
Matches 38; Conservative 22; Mismatches 29; Indels 15; Gaps 3;
Qy 10 GGGGPAKRAAGD-----GPSESADD--DIVVAQISKNNRVAVRTWNGKV 55
Db 76 GGAGEGKDEQAESEEEEEERGGGTYKRYDDGLILCLARRVTLQEFKGT 135
Qy 56 VVDIREFYKDGKTLPGKGIQLPMDQWKILRDNIKAIDEAIKE 99
Db 136 LLSIREYFKDGKELPA-KGISLTVEQWEAFRDSVPAIEDAIKK 178
RESULT 7
TCP4 MOUSE
ID TCP4 MOUSE STANDARD; PRT; 126 AA.
AC P11031;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Activated RNA polymerase II transcriptional coactivator p15 precursor
DE (Positive cofactor 4) (PC4) (p14) (Single-stranded DNA binding protein
DE p9).
GN Name=Rpo2tcl; Synonym=PC4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88228080; PubMed=3372536;
RA Ballard D.W., Philbrick W.M., Bothwell A.L.M.;
RT "Identification of a novel 9-kDa polypeptide from nuclear extracts.
RT DNA binding properties, primary structure, and in vitro expression.";
RL J. Biol. Chem. 263:8450-8457(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Czech II; TISSUE=Mammary gland;
RX MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Munzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schenck A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -I- FUNCTION: General coactivator that functions cooperatively with
CC TAFs and mediates functional interactions between upstream
CC activators and the general transcriptional machinery. Binds
CC single-stranded DNA.
CC -I- SUBCELLULAR LOCATION: Nuclear.
CC -I- PTM: Proteolysis is required for p9 DNA-binding activity.
CC -I- PTM: Activity of p15 is controlled by protein kinases that target
CC the regulatory domain. Phosphorylation inactivates cofactor
CC function (By similarity).
CC -I- SIMILARITY: Belongs to the transcriptional coactivator PC4 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```





```
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative transcriptional co-activator (At5g09240).
GN Name=TSB8_40; Synonyms=At5g09240/TSB8_40;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
RA Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
RA Hayashizaki Y., Shinozaki K.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL391712; CAC05450.1; -
DR EMBL; AK117904; BAC42542.1; -
DR EMBL; BT004696; AAO42942.1; -
DR HSSP; P53999; 1PCF.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003713; F:transcription coactivator activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003173; PC4.
DR InterPro; IPR003044; sDNA_bind_regul.
DR Pfam; PF02229; PC4; 1.
SQ SEQUENCE 110 AA; 12699 MW; C935C3BB613F949D CRC64;

Query Match 31.1%; Score 165; DB 2; Length 110;
Best Local Similarity 40.4%; Pred. No. 1.1e-08;
Matches 38; Conservative 19; Mismatches 23; Indels 14; Gaps 3;

Qy 20 RAAGDDGPSES-----AD--DDIVVAQISKRRVAVRTWNGKVVVDIREFYBKDG 67
Db 13 RASDRDESETHAPPKVAKPADEIDFICNLQKRRVFRVNCNGRIWIAIRFFVKDG 72

Qy 68 KTLF--GRKGQLPMDQWIKLRDNIKAIDEAIKE 99
Db 73 ITLPCKSQGISLSLEQWDLNRHEEDIKALSE 106

RESULT 12
ID Q6E433 PRELIMINARY; PRT; 73 AA.
AC Q6E433;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Activated RNA polymerase II transcription cofactor 4 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Simoes-Barbosa A., Teixeira A.R.L.;
```

```
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY584192; AAT72821.1; -
FT NON TER 1
SQ SEQUENCE 73 AA; 8655 MW; E566BF30D7B781CD CRC64;

Query Match 30.5%; Score 161.5; DB 2; Length 73;
Best Local Similarity 45.1%; Pred. No. 1.6e-08;
Matches 32; Conservative 19; Mismatches 17; Indels 3; Gaps 2;

Qy 30 SADDIVVAQISKRRVAVRTWNGKVVVDIREFY-ERDGTLPGRKGITQLPMDQWIKLRD 88
Db 4 SRDDNMF--QIGKRWYVSVDRDFPKKILIDIREYMDSEGEKPKGRKISLNMEQWSQLKE 61
Qy 89 NIKAIIDEAIKE 99
Db 62 QISDIDDAVRK 72

RESULT 13
TCP4 HUMAN
ID TCP4 HUMAN STANDARD; PRT; 126 AA.
AC P53999; Q96L29;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Activated RNA polymerase II transcriptional coactivator p15 (Positive cofactor 4) (PC4) (p14).
GN Name=RPO2TC1; Synonyms=PC4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX MEDLINE=94340741; PubMed=8062392;
RA Kretschmar M., Kaiser K., Lottspeich F., Meisterernst M.;
RT "A novel mediator of class II gene transcription with homology to viral immediate-early transcriptional regulators.";
RL Cell 78:525-534 (1994).
RN [2]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 1-10 AND 80-96.
RX MEDLINE=94340740; PubMed=8062391;
RA Ge H., Roeder R.G.;
RT "Purification, cloning, and characterization of a human coactivator, PC4, that mediates transcriptional activation of class II genes.";
RL Cell 78:513-523 (1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow, Cervix, Lung, and Prostate;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haiech F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Sapletenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A.C., Shevchenko Y., Bouffard G.G.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Butterfield J.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [4]
RP PHOSPHORYLATION, AND MASS SPECTROMETRY.
RX MEDLINE=95108024; PubMed=7809103;
```



Job time : 196 secs

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**



Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	117	22.1	292	4	US-09-538-092-608	Sequence 608, Appl	
2	106	20.6	159	4	US-09-448-796A-19133	Sequence 19133, A	
3	82.5	15.6	11877	3	US-09-205-537-6	Sequence 6, Appl	
4	75	14.2	736	4	US-09-252-991A-31194	Sequence 31194, A	
5	74	14.0	275	4	US-09-543-681A-7448	Sequence 7448, Ap	
6	72	13.6	237	4	US-09-252-991A-17245	Sequence 17245, A	
7	70.5	13.3	159	3	US-08-990-791-7	Sequence 7, Appl	
8	70.5	13.3	159	4	US-09-372-591-7	Sequence 7, Appl	
9	70	13.2	1404	3	US-08-801-308-1	Sequence 1, Appl	
10	68.5	12.9	434	4	US-09-594-193-7	Sequence 7, Appl	
11	68	12.8	525	2	US-08-676-166A-3	Sequence 3, Appl	
12	67	12.6	495	4	US-09-252-991A-22825	Sequence 22825, A	
13	66.5	12.5	416	4	US-09-270-767-43018	Sequence 43018, A	
14	66	12.5	154	4	US-09-548-938A-19	Sequence 19, Appl	
15	66	12.5	172	4	US-09-270-767-32192	Sequence 32192, A	
16	66	12.5	172	4	US-09-270-767-47409	Sequence 47409, A	
17	66	12.5	526	4	US-09-548-938A-10	Sequence 10, Appl	
18	66	12.5	1016	4	US-09-635-972-24	Sequence 24, Appl	
19	65.5	12.4	159	3	US-08-990-791-6	Sequence 6, Appl	
20	65.5	12.4	159	4	US-09-372-591-6	Sequence 6, Appl	
21	65.5	12.4	160	3	US-08-809-326A-14	Sequence 14, Appl	
22	65.5	12.4	160	4	US-09-689-914A-14	Sequence 14, Appl	
23	65.5	12.4	160	4	US-09-689-913A-14	Sequence 14, Appl	
24	65.5	12.4	160	4	US-09-689-916A-14	Sequence 14, Appl	
25	65.5	12.4	432	3	US-08-809-326A-16	Sequence 16, Appl	
26	65.5	12.4	432	4	US-09-689-914A-16	Sequence 16, Appl	
27	65.5	12.4	432	4	US-09-689-913A-16	Sequence 16, Appl	

; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 19133  
; LENGTH: 159  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-19133

Query Match 20.0%; Score 106; DB 4; Length 159;  
Best Local Similarity 34.1%; Pred. No. 0.00011;  
Matches 29; Conservative 18; Mismatches 30; Indels 8; Gaps 3;  
QY 16 AKRRAAGDDGPSESADDDIVVAQISKRRVAVRTWNGKVVVDIREFY-EKDGKTLPGKK 74  
DB 15 AFRGA-----PASNSNEVEILLDLKK--QVTVKYNVNLVDIREFYTRDGVKPGKK 67  
QY 75 GLOPMDQWILRDNIKAIDEAIKE 99  
DB 68 GISLTEDTYKKLEATNKIQNALDD 92

RESULT 3

US-09-105-537-6  
; Sequence 6, Application US/09105537A  
; Patent No. 6265202  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, D.H.  
; APPLICANT: Liu, H.  
; APPLICANT: Xue, Y.  
; APPLICANT: Zhao, L.  
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
; FILE REFERENCE: 600.438US1  
; CURRENT APPLICATION NUMBER: US/09/105,537A  
; CURRENT FILING DATE: 1998-06-26  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 6  
; LENGTH: 11877  
; TYPE: PRT  
; ORGANISM: Streptomyces venezuelae  
US-09-105-537-6

Query Match 15.6%; Score 82.5; DB 3; Length 11877;  
Best Local Similarity 26.9%; Pred. No. 14;  
Matches 32; Conservative 17; Mismatches 35; Indels 35; Gaps 6;  
QY 10 GGGGEPAAKER-----AAGDDGP-----SESADDDIVV-----AQISKRRVAVRTWN 52  
DB 8641 GGRPELAARLRAALAAALGGDDGATDLDEASDDDLFSFDKELGSDSDFMANNEDKRDYL 8700  
QY 53 GKVVVDI-----REFYKDGKT-----LPGRKGIGLPMDQWILRDNIKAIDE 95  
DB 8701 KRVTAELQNTRLREIEGRTHFPAIVGVNACRLPG--GVASPEDLWQVAGDGAISE 8757

RESULT 4

US-09-252-991A-31194  
; Sequence 31194, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; PRIOR FILING DATE: 1998-02-18  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 31194  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-31194

Query Match 14.2%; Score 75; DB 4; Length 736;  
Best Local Similarity 34.7%; Pred. No. 3.2;  
Matches 26; Conservative 9; Mismatches 20; Indels 20; Gaps 4;  
QY 2 WRKGNKRFGG---GGSE-----PAKRRAAGDDGPSESADDDIVVAQISKRRVAVRTWN 52  
DB 636 FRGRLRFGGGAGGDRDAEGBEGERRAAGQDLAAAEAGDQVVHVVRV-----VRMVA 688  
QY 53 GKVVVDIREFYEKDG 67  
DB 689 GOLVA-----ITEQDG 699

RESULT 5

US-09-543-681A-7448  
; Sequence 7448, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 7448  
; LENGTH: 275  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-7448

Query Match 14.0%; Score 74; DB 4; Length 275;  
Best Local Similarity 29.5%; Pred. No. 1.2;  
Matches 18; Conservative 14; Mismatches 23; Indels 6; Gaps 2;  
QY 39 QISKRRVAVRTW--NGKVVVDIREFYEKD-----GKTLPGRKGIGLPMDQWILRDNIKA 92  
DB 146 QSKRKKIAIKPWLMDKLVGVGVNIYANEALFSSGIMPRKANSLTEQCDVIVNAIKA 205  
QY 93 I 93  
DB 206 V 206

RESULT 6

US-09-252-991A-17245  
; Sequence 17245, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; PRIOR FILING DATE: 1998-02-18  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 17245  
; LENGTH: 237  
; TYPE: PRT

2GRK 74

RESULT 11  
US-08-676-166A-3  
; Sequence 3, Application US/08676166A  
; Patent No. 5955270  
; GENERAL INFORMATION:  
; APPLICANT: Radford, Alan  
; APPLICANT: Parish, John H.  
; TITLE OF INVENTION: EXPLOITATION OF THE CELLULASE COMPLEX OF  
; TITLE OF INVENTION: NEUROSPORA  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David A. Jackson, Esq.  
; STREET: 411 Hackensack Ave, Continental Plaza, 4th  
; STREET: Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/676,166A  
; FILING DATE: 15-JUL-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742



THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2005, 23:06:53 ; Search time 144 Seconds  
(without alignments)  
252.901 Million cell updates/sec

Title: US-10-629-953-4

Perfect score: 530

Sequence: 1 MWRKGNKRFGGGEPAAKRR.....QWKILRDNIKAIDEAIKENA 101

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1603904 seqs, 360571292 residues

Total number of hits satisfying chosen parameters: 1603904

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	530	100.0	475	16	US-10-437-963-185855
2	398	75.1	119	15	US-10-425-114-42612
3	395	74.5	103	17	US-10-425-115-259005
4	269	50.8	105	15	US-10-424-599-181609
5	261.5	49.3	105	15	US-10-424-599-181610
6	182.5	34.4	163	15	US-10-424-599-147198
7	182.5	34.4	196	15	US-10-425-114-38036
8	180.5	34.1	184	16	US-10-437-963-162084
9	178.5	33.7	367	15	US-10-424-599-151999
10	177	33.4	79	15	US-10-424-599-270637
11	172	32.5	141	13	US-10-087-192-1785
12	165	31.1	140	17	US-10-425-115-205102
13	165	31.1	150	15	US-10-425-114-67613

14	165	31.1	183	17	US-10-425-115-339084
15	165	31.1	198	15	US-10-425-114-60171
16	165	31.1	204	15	US-10-425-114-69791
17	160.5	30.3	127	13	US-10-087-192-1788
18	160.5	30.3	127	14	US-10-177-293-350
19	160.5	30.3	127	14	US-10-257-021-22
20	160.5	30.3	127	17	US-10-370-7158-282
21	160.5	30.3	132	9	US-09-925-297-779
22	153.5	29.0	106	16	US-10-767-701-60649
23	110	20.8	104	11	US-09-864-408A-4696
24	107.5	20.3	82	17	US-10-425-115-342501
25	101.5	19.2	88	17	US-10-425-115-251977
26	87	16.4	124	17	US-10-425-115-202622
27	83.5	15.6	87	17	US-10-425-115-272770
28	82.5	15.6	11877	9	US-09-861-289-6
29	82.5	15.6	11877	9	US-09-860-846-6
30	82.5	15.6	11877	10	US-09-836-821-6
31	82.5	15.6	11877	14	US-10-271-889-49
32	82.5	15.6	12159	10	US-09-988-3848-6
33	80	15.1	118	17	US-10-425-115-334183
34	79.5	15.0	115	16	US-10-767-701-34940
35	79.5	15.0	134	17	US-10-425-115-286086
36	79.5	15.0	146	17	US-10-425-115-341160
37	76.5	14.4	189	16	US-10-437-963-104003
38	76	14.3	507	16	US-10-437-963-127962
39	76	14.3	924	17	US-10-425-115-358126
40	75.5	14.2	499	15	US-10-282-122A-61693
41	74.5	14.1	100	16	US-10-767-701-37709
42	74.5	14.1	465	15	US-10-282-122A-47419
43	74	14.0	506	16	US-10-437-963-164624
44	73.5	13.9	128	17	US-10-425-115-336013
45	73	13.8	263	16	US-10-325-468-26

ALIGNMENTS

RESULT 1

US-10-437-963-185855  
; Sequence 185855, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 185855  
; LENGTH: 475  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_8270C.1.pep  
US-10-437-963-185855

Query Match 100.0%; Score 530; DB 16; Length 475;  
Best local Similarity 100.0%; Pred. No. 8.5e-52;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWRKGNKRFGGGEPAAKRRAGDGPSEADDDIVVAQISKNNRVAVRTWNGKVVDIR 60

Db 375 MWRKGNKRFGGGEPAAKRRAGDGPSEADDDIVVAQISKNNRVAVRTWNGKVVDIR 434

QY 61 EFYKDGKTLPGRGIGIQLPMDQWKILRDNIKAIDEAIKENA 101







Db 86 KKENEDIAAKNDVAKKEDVVVANEBSRETEVLPLKLRDDPVRVCHLSNRRNVAVKDF 145  
Qy 52 NGKVVVDIREFYKDGKTLPRGKIGQLPMDQWILRDNKAIDEAIKE 99  
Db 146 KGTTLVSIREFYKDGKPLPGSKGILSLSSQWSTFKKSVPAIEAIKK 193

## RESULT 10

US-10-424-599-270637  
; Sequence 270637, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 270637  
; LENGTH: 79  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_86402C.1.pap  
US-10-424-599-270637

Query Match 33.4%; Score 177; DB 15; Length 79;  
Best Local Similarity 41.9%; Pred. No. 2-2e-12;  
Matches 31; Conservative 18; Mismatches 25; Indels 0; Gaps 0;

Qy 26 GPSESADDDIVVAQISKNNRRVAVRTWNGKVVVDIREFYKDGKTLPRGKIGQLPMDQWKI 85  
Db 2 GKEYDDGDLIICRLSKRRVTIQDFRGKTLVSIREFYKDGKTLPRGKIGQLPMDQWKI 61  
Qy 86 LRDNKAIDEAIKE 99  
Db 62 FKQVPAIEAIKK 75

## RESULT 11

US-10-087-192-1785  
; Sequence 1785, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1785  
; LENGTH: 141  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_86402C.1.pap  
US-10-087-192-1785

Query Match 32.5%; Score 172; DB 13; Length 141;  
Best Local Similarity 41.4%; Pred. No. 1.7e-11;  
Matches 36; Conservative 20; Mismatches 29; Indels 2; Gaps 2;

Qy 14 EPAKGRAGDGPSESADDDIVVAQISKNNRRVAVRTWNGKVVVDIREFY-EKDGKTLPG 72

Db 55 KPGETSRALASSQSSSSRDDNMF-QIGKRVYVSRDFKGLIDIREYVMDSEGMKPG 113  
Qy 73 RKGIQLPMDQWILRDNKAIDEAIKE 99  
Db 114 RKGISLNBQWSQLKEQISDIDDAVRK 140

## RESULT 12

US-10-425-115-205102  
; Sequence 205102, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 205102  
; LENGTH: 140  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_118642C.1.pap  
US-10-425-115-205102

Query Match 31.1%; Score 165; DB 17; Length 140;  
Best Local Similarity 40.5%; Pred. No. 1.1e-10;  
Matches 32; Conservative 17; Mismatches 28; Indels 2; Gaps 1;

Qy 23 GDDGPSESADD-DIVVAQISKNNRRVAVRTWNGKVVVDIREFYKDGKTLPRGKIGQLPM 80  
Db 57 GKGGKRGYVDDQDGLILCRSLSKRRVTLSEFKGRSLVIREFYKDGKEMPSAKGISMTM 116  
Qy 81 DQWKILRDNKAIDEAIKE 99  
Db 117 EQWEAFNCNAVPAIEAIKK 135

## RESULT 13

US-10-425-114-67613  
; Sequence 67613, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 67613  
; LENGTH: 150  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3279-041-H1\_FLI.pap  
US-10-425-114-67613

Query Match 31.1%; Score 165; DB 15; Length 150;  
Best Local Similarity 40.5%; Pred. No. 1.1e-10;  
Matches 32; Conservative 17; Mismatches 28; Indels 2; Gaps 1;

Qy 23 GDDGPSAD--DIVVAQISKRRVAVRTWNGKVVVDIREFYKDGKTLPGKGIOLPM 80  
Db 67 GKGGKREYDDQGLILCRLLSKRRVTLSEFKGRSLVIREFYVKDGKEMPSAKGISMTM 126  
Qy 81 DOWKILRDNIKAIDEAIKE 99  
Db 127 EQWEAFNCNVAPEAIAIKK 145

RESULT 14  
US-10-425-115-339084  
; Sequence 339084, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 339084  
; LENGTH: 183  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_72412C.1.pap  
US-10-425-115-339084

Query Match 31.1%; Score 165; DB 17; Length 183;  
Best Local Similarity 39.2%; Pred. No. 1.5e-10;  
Matches 31; Conservative 19; Mismatches 27; Indels 2; Gaps 1;

Qy 23 GDDGPSAD--DIVVAQISKRRVAVRTWNGKVVVDIREFYKDGKTLPGKGIOLPM 80  
Db 100 GKGGKREYDDQGLILCRLLSKRRVTLSEFKGRSLVIREFYVKDGKEMPSAKGISMTL 159  
Qy 81 DOWKILRDNIKAIDEAIKE 99  
Db 160 EQWEAFNCNVAPEAIAIKK 178

RESULT 15  
US-10-425-114-60171  
; Sequence 60171, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 60171  
; LENGTH: 198  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3608-031-G8\_FLI.pap  
US-10-425-114-60171

Query Match 31.1%; Score 165; DB 15; Length 198;  
Best Local Similarity 39.2%; Pred. No. 1.6e-10;  
Matches 31; Conservative 19; Mismatches 27; Indels 2; Gaps 1;

Qy 23 GDDGPSAD--DIVVAQISKRRVAVRTWNGKVVVDIREFYKDGKTLPGKGIOLPM 80  
Db 115 GKGGKREYDDQGLILCRLLSKRRVTLSEFKGRSLVIREFYVKDGKEMPSAKGISMTL 174  
Qy 81 DOWKILRDNIKAIDEAIKE 99  
Db 175 EQWEAFNCNVAPEAIAIKK 193

Search completed: January 10, 2005, 23:19:10  
Job time : 145 secs

THIS PAGE BLANK (uspto)

THIS PAGE BLANK (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 11, 2005, 23:15:54 ; Search time 3302 Seconds  
(without alignments)  
1446.475 Million cell updates/sec

Title: US-10-629-953-4

Perfect score: 530

Sequence: 1 MWRKGNRRGGGEPAAKRR.....QWKILRDNIKAIDEAIKENA 101

Scoring table: BLOSUM62

-Q= /cpn2 1/USPTO.spool.h/US10629953/runat 10012005 172535 9414/app query fasta\_1.263  
-DB=GenEmbl -QFWT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10629953 @CGN 1.1 3731 @runat 10012005 172535 9414 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DSLEX=7

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q= /cpn2 1/USPTO.spool.h/US10629953/runat 10012005 172535 9414/app query fasta\_1.263  
-DB=GenEmbl -QFWT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10629953 @CGN 1.1 3731 @runat 10012005 172535 9414 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DSLEX=7

Database :

GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	530	100.0	691	8 AK072440	AK072440 Oryza sat
2	368	69.4	289	6 AR246100	AR246100 Sequence
3	352	66.4	145419	8 AP004876	AP004876 Oryza sat
4	297	56.0	263	6 AR246253	AR246253 Sequence

#### ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
AK072440	Oryza sativa (japonica cultivar-group)	691 bp mRNA linear	AK072440	AK072440.1	GI:32982463	FLI CDNA: CAP trapper.	Oryza sativa (japonica cultivar-group)	1	The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team, Kikuchi S., Satoh K., Nagata T., Kawagashira N., Doi K., Kishimoto N., Yazaki J., Ishikawa M., Yamada H., Ooka H., Hotta I., Kojima K., Namiki T., Ohneda E., Yahagi W., Suzuki K., Li C., Ohtsuki K., Shishiki T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Otonari Y., Murakami K., Kudo Y., Sugano S., Fujimura T., Suzuki Y., Tsunoda Y., Kurokawa R., Sugiyama A., Mizuno K., Yokomizo S., Miura J., Ikeda R., Ishibiki J., Kawamata M., Yoshimura A., Miura J., Kusumegi T., Oka M., Ryu R., Ueda M., Matsubara K., RIKEN, Kawai J., Carninci P., Adachi J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashidume W., Hayatsu N., Imotani K., Ishii Y., Itoh M.,
AK072440	Oryza sativa (japonica cultivar-group)	691 bp mRNA linear	AK072440	AK072440.1	GI:32982463	FLI CDNA: CAP trapper.	Oryza sativa (japonica cultivar-group)	1	The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team, Kikuchi S., Satoh K., Nagata T., Kawagashira N., Doi K., Kishimoto N., Yazaki J., Ishikawa M., Yamada H., Ooka H., Hotta I., Kojima K., Namiki T., Ohneda E., Yahagi W., Suzuki K., Li C., Ohtsuki K., Shishiki T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Otonari Y., Murakami K., Kudo Y., Sugano S., Fujimura T., Suzuki Y., Tsunoda Y., Kurokawa R., Sugiyama A., Mizuno K., Yokomizo S., Miura J., Ikeda R., Ishibiki J., Kawamata M., Yoshimura A., Miura J., Kusumegi T., Oka M., Ryu R., Ueda M., Matsubara K., RIKEN, Kawai J., Carninci P., Adachi J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashidume W., Hayatsu N., Imotani K., Ishii Y., Itoh M.,





2



Db 186 TTACGGAGTCTTACTTCAAGGACGGCAAGACTCTCCCAACCGCAAGGTATATC-ATC 244

Qy 79 ProMetAspGlnTrpLys 84  
:::|||||

Db 245 CAATTAGATCAGTGAAG 262

RESULT 5

AR251557

LOCUS AR251557 286 bp DNA linear PAT 20-DEC-2002

DEFINITION .Sequence 6916 from patent US 6476212.

ACCESSION AR251557

VERSION AR251557.1 GI:27299431

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 286)

AUTHORS Lalugudi,R.V., Ito,L.Y. and Sherman,B.K.

TITLE Polynucleotides and polypeptides derived from corn ear

JOURNAL Patent: US 6476212-A 6916 05-NOV-2002;

FEATURES

source

1..286

/organism="unknown"

/mol\_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 9.35e-18 Length: 286

Score: 275.00 Matches: 57

Percent Similarity: 83.54% Conservative: 9

Best Local Similarity: 72.15% Mismatches: 10

Query Match: 51.89% Indels: 3

DB: 6 Gaps: 1

US-10-629-953-4 (1-101) x AR251557 (1-286)

Qy 1 MetTrpArgLysGlyAenLysArgPhe-GlyGlyGly-----GlyGluProAlaLalY 18  
|||||

Db 50 ATGTNGGGGAAGAAAGAACGCTTCGGCGGTGGCGGAGCGGTGAGCCGCGGCCAA 109  
|||||

Qy 18 sArgArgAlaAlaGlyAspArgGlyProSerGluSerAlaAspAspLysValValAl 38  
|||||

Db 110 GCGCCAGCGCGGAGGACGATGCGCTCCGATCCGCTGAAGACGGTACCGTCTCGTAGC 169  
|||||

Qy 38 aGlnIleSerLysAsnArgValAlaValAlaArgThrTrpAsnGlyLysValValValas 58  
|||||

Db 170 CGAGATATCGACACAAAGAGGTGTCCTAGGAGCTGGAAGGACGGTCTTCGTGCA 229  
|||||

Qy 58 pLleArgGluPheTrpGluLysAspGlyLysThrLeuProGlyArgLysGlylle 76  
|||||

Db 230 CTTACGCGAGTCTTACTTCAAGGACGGCAAGACTCTCCCAACCGCAAGGTATA 284  
|||||

RESULT 6

BT004745

LOCUS BT004745 324 bp mRNA linear PLN 22-FEB-2003

DEFINITION Arabidopsis thaliana At5g09250 gene, complete cds.

ACCESSION BT004745

VERSION BT004745.1 GI:28466804

KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 324)

AUTHORS Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Bowser,L., Carninci,P., Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Huan,V.W., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.

TITLE Arabidopsis ORF clones

JOURNAL Unpublished

AUTHORS 2 (bases 1 to 324)

Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Bowser,L., Carninci,P., Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Huan,V.W., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.

TITLE Direct Submission

JOURNAL Submitted (22-FEB-2003) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL CDNA: 'RIKEN Arabidopsis Full-length CDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGECC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Bowser,L., Chan,M.M., Chang,C.H., Dale,J.M., Huan,V.W., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Davis,R.W., Theologis,A., and Ecker,J.R.

Kim,C.J. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

source

1..324

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

-/db\_xref="taxon:3702"

/chromosome="5"

/clone="U23341"

/ecotype="Columbia"

/notes="This clone is in pUNI 51"

CDS

1..324

/notes="putative transcriptional co-activator (KIWI)"

/codon\_start=1

/product="At5g09250"

/protein\_id="AA044011.1"

/translation="MSSRGKRKDEVDVRSDESETHAPAKKVAKPADDSDSDIVVC NISKNRVSVRNWNGKIWIIDREFYVKGKTLFGKIGISLSDQWNTLRNHAEDIEKA LSDLS"

ORIGIN

Alignment Scores:

Pred. No.: 4.43e-15 Length: 324

Score: 247.50 Matches: 48

Percent Similarity: 70.59% Conservative: 12

Best Local Similarity: 56.47% Mismatches: 22

Query Match: 46.70% Indels: 3

DB: 8 Gaps: 1

US-10-629-953-4 (1-101) x BT004745 (1-324)

Qy 15 ProAlaAlaLysArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAsp 34  
|||||

Db 70 CCGGCTAAGAAAGTCCGAAAGCCAGCAGCAT-----TCCGACCATCTTGACGAT 120  
|||||

Qy 35 IleValValAlaGlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLys 54  
|||||

Db 121 ATGTCGTCTGCAACATATCTTAGAATAGGAGATCTCTGTAGGAATTCGACGGGAAG 180  
|||||

Qy 55 ValValValAspLysArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLys 74  
|||||





modified two-hybrid system incorporating a GFP reporter gene

Plant J. 14 (6), 685-692 (1998)

98346011

9681033

2 (bases 1 to 730)

Cormack, R.S., Hahnbrock, K. and Somssich, I.E.

Direct Submission

Submitted (11-MAR-1998) Biochemistry, Max-Planck-Institut fuer

Zuechtungsforchung, Carl-von-Linne-Weg 10, Cologne 50829, Germany

Location/Qualifiers

1. .730

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/cultivat="Columbia"

/db\_xref="taxon:3702"

/chromosomes="IV"

/map="between m518 and BIO206"

1. .730

/genes="KELP"

45. .542

/gene="KELP"

/codon\_start=1

/product="putative transcriptional co-activator"

/protein\_id="AAC08575.1"

/db\_xref="GI:2997686"

/translation="MEKETKEIKETVIELSEDMKEITEFKVKLASEKLAIDLSE  
KSHKAFVRSVVEKFLDEERAREYENSOVNKEEDGDKCGKNGKEDDDGLICRLS  
DKRRVTIQEFKGSLSVIREYKDKGKELPTSGISLTDEQWSTFKKNPAENAVKK  
MESRV"

ORIGIN

Alignment Scores:

Pred. No.: 1.45e-08 Length: 730

Score: 183.00 Matches: 37

Percent Similarity: 56.25% Conservative: 17

Best Local Similarity: 38.54% Mismatches: 22

Query Match: 34.53% Indels: 20

DB: 8 Gaps: 2

US-10-629-953-4 (1-101) x AF053303 (1-730)

QY 4 LyeGlyAsnLysArgPheGlyGlyGlyGluProAlaAlaLysArgAlaAlaGly 23

Db 297 AAAGGAACAAGAGTTT----- 314

QY 24 AspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLysAsn 43

Db 315 GATGAT-----GACGGCGATCTTATCATTTGCGAGTTATCGGATAAG 356

QY 44 ArgArgValAlaValArgThrTrpAenGlyLysValValValAspIleArgGluPheTyr 63

Db 357 AGAAGAGTGACGANTCAGGAANTTAAGGGAGAGTTTGGTTCTTATCAGAGTATTAC 416

QY 64 GLuLYeAspGlyLysThrLeuProGlyArgLysGlyLeuGlnLeuProMetAspGlnTrp 83

Db 417 AAGAAAGATGCAAGAACTTCTACTCTTAAGGAATAAGCTTAACAGATGAACAATGG 476

QY 84 LysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99

Db 477 TCAACCTTCAAGAAACATGCCAGCCATCGAAANTGCTGTCAAGAA 524

RESULT 11

AY085896

LOCUS

DEFINITION

Arabidopsis thaliana clone 19449 mRNA, complete sequence.

ACCESSION

AY085896

VERSION

AY085896.1 GI:21404606

KEYWORDS

FLI CDNA.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

REFERENCE

AUTHORS

1 (bases 1 to 757)

Haas, B.J., Volfovsky, N., Town, C.D., Troukhan, M., Alexandrov, N.,  
Feldmann, K.A., Flavell, R.B., White, O. and Salzberg, S.L.

Full-length messenger RNA sequences greatly improve genome  
annotation

Genome Biol. 3 (6), RESEARCH0029 (2002)

22088475

12093376

2 (bases 1 to 757)

Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and  
Feldmann, K.

Full-length cDNA from Arabidopsis thaliana  
Unpublished

3 (bases 1 to 757)

Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and  
Feldmann, K.

Direct Submission

Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,  
Malibu, CA 90265, USA

This clone sequence is one of 5,000 Ceres full-length cDNAs made  
available to TIGR and Genbank. The following quality assessment of  
this set was done by comparison with known proteins: two percent of  
the clones are estimated to be 5'-truncated; less than one percent  
are 3'-truncated; approximately two percent represent alternative  
splice variants, including unspliced introns and spliced exons; one  
percent may contain premature stop codons; five percent may have  
frame shifts in a coding region. A sequence is considered to be  
5'-truncated if it lacks the translation initiation start (ATG). A  
sequence is considered to be 3'-truncated if it lacks the  
C-terminal end of the encoded protein. Please note that these cDNA  
sequences are derived from the WS or Laer ecotypes and therefore  
may contain polymorphisms when compared to sequences from Col-0.  
Genset carried out the library production and sequencing of the  
full-length clones. Ceres, Inc. carried out the clustering of the  
5' sequences, selection of clones, and sequence assembly.

FEATURES

source

1. .757

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/clone="19449"

91. .588

/codon\_start=1

/product="putative transcriptional coactivator"

/protein\_id="AAM63108.1"

/db\_xref="GI:21554027"

/translation="MEKETKEIKETVIELSEDMKEITEFKVKLASEKLAIDLSE  
KSHKAFVRSVVEKFLDEERAREYENSOVNKEEDGDKCGKNGKEDDDGLICRLS  
DKRRVTIQEFKGSLSVIREYKDKGKELPTSGISLTDEQWSTFKKNPAENAVKK  
MESRV"

ORIGIN

Alignment Scores:

Pred. No.: 1.51e-08 Length: 757

Score: 183.00 Matches: 37

Percent Similarity: 56.25% Conservative: 17

Best Local Similarity: 38.54% Mismatches: 22

Query Match: 34.53% Indels: 20

DB: 8 Gaps: 2

US-10-629-953-4 (1-101) x AY085896 (1-757)

QY 4 LysGlyAsnLysArgPheGlyGlyGlyGluProAlaAlaLysArgAlaAlaGly 23

Db 343 AAAGGAACAAGAGTTT----- 360

QY 24 AspAspGlyProSerGluSerAlaAspAspIleValValAlaGlnIleSerLysAsn 43

Db 361 GATGAT-----GACGGCGATCTTATCATTTGCGAGTTATCGGATAAG 402

QY 44 ArgArgValAlaValArgThrTrpAenGlyLysValValValAspIleArgGluPheTyr 63

Db 403 AGAAGAGTACGANTCAGGAANTTAAGGGAGAGTTTGGTTCTTATCAGAGTATTAC 462









gene	/notes="overlap to PAC MAH20, EMBL:AB006697; please refer to this clone for annotation"	exon	/gene="At5g08660"
CDS	15981..19300	intron	/number=8
	/gene="At5g08660"		18231..18347
	join(15981..16289,16433..16746,16833..16925,17020..17110,17305..17478,17575,17767..17836,18028..18131,18231..18347,18458..18534,18636..18680,18763..19300)		/gene="At5g08660"
	/gene="At5g08660"	exon	/number=9
	thaliana"		18348..18457
	/codon_start=1		/gene="At5g08660"
	/product="putative protein"	intron	/number=9
	/protein_id="CAC35871.1"		18458..18534
	/db_xref="GI:13548324"		/gene="At5g08660"
	/db_xref="TrEMBL:O9C5B1"	exon	/number=10
	/translation="MGSFCSKSLGINSFSGSVADGREDPFGHSOPNGOTSLIV		18535..18635
	PGNRQLMKVDKQKQQLKDVFSFRERAEADNPYDGIPTVTMAPSQKRSKQTQTAVI	intron	/gene="At5g08660"
	KVTEASLKGKAGKRAKVDLTIGSSMTDLGSGFTSGVATKGNELGILAFEVANT		/number=10
	VKSNTJESLSKRNIEHLKGTILYSEGVQLVSNDFDLRLVAADKQELQVFSGEV	exon	18636..18680
	VRFGNRSKDFQWNLQRYFDRI SKELTPQRLKEDAVLVVDQLMVLVQVTAELYQELQ		/gene="At5g08660"
	VLVRLKDYEKREERENSANSKGDGLAILTELKAQKVVKSLKKSLWRSRFEV	intron	/number=11
	MEKLVDIVHELLEIHNIPGADDPQSKGAAYDKRLGPAGLALHYANILVQIDTLV		18681..18762
	ARASSITSNARDSLYSLPGLKALRSKISFNVDKSLSVQIKDEMERTLHLVFPV	exon	/gene="At5g08660"
	AGNTVKAHGFVGWGWANTGTDTFSKSGDILRIETLYHASKETEIYILGQIILWL		/number=11
	QHLVTAKSDARGPLRSITKSPDTPTNQQLISEPLSVPTVDEQKMLQEASKRRT	gene	/gene="At5g08660"
	PCVSKSQDFSEYSRARKCDPLSKSEYFRGVRRSKSAVKRYSSGFFLLDFAIDKEK	gene	18763..19300
	VLVDIVRDVPRDYKALLKEGSLSF"		/number=12
exon	15981..16289		/gene="At5g08660"
	/gene="At5g08660"		1950..22604
intron	/number=1	CDS	/gene="At5g08670"
	16290..16432		complement(join(19850..19963,20204..20515,20610..20834,20918..21076,21177..21280,21360..21477,21569..21657,21738..21813,22131..22604))
	/gene="At5g08660"		/gene="At5g08670"
exon	16433..16746		/notes="strong similarity to H+-transporting ATP synthase, mitochondrial, rubber tree, PIR:S20504
intron	/gene="At5g08660"		Contains ATP/GTP-binding site motif A (P-loop)
	/number=2		AA231-238;ATP synthase alpha and beta subunits signature
	16747..16832		AA423-432"
	/gene="At5g08660"		/codon_start=1
exon	/number=2		/product="H+-transporting ATP synthase beta chain
	16833..16925		(mitochondrial)-like protein"
intron	/gene="At5g08660"		/protein_id="CAC35872.1"
	/number=3		/db_xref="GI:13548325"
	16926..17019		/db_xref="GOA:P83483"
	/gene="At5g08660"		/db_xref="Swiss-Prot:P83483"
exon	/number=3		/translation="WASREVLSSLRSSRGAALKGNRPRLPSPSPARHAAPCSVL
intron	17020..17110		LCRVAYATSSPASSAPSSAPAKBEGKTYDYGKGAIGRVQVIGALVDVFEQOE
	/gene="At5g08660"		GLPPIMTSLEVDQHPTRLVLEVSHHLGQNVVRIAMDGTEGLVRGRKVLNTGPIVTP
	/number=4		VGRATLGRIMNVLGEFIDBERGKTEHYLPPIHRDAPALVDLATGQELATGKVVDLL
	17111..17304		APYQGGKIGLFGAGVGKTVLMELINNVAKAHGGSFVAGVGTREGNDLYREMI
	/gene="At5g08660"		ESGVIKLGKQSKCALVQWNEPPGARARVGLTVAEYFRDAEGQDVLIFDN
exon	/number=4		IFRPTQANSEVSALLGRIPSAVGYOPTLADLGALEOERTTTTKGGSITSVOALVVPAD
intron	17305..17398		VDTPAPATTFALHDATTVLSRQISELGIYPADVPLDSTSRMLSPHILGEHYNTARG
	/gene="At5g08660"		VLKVNQYNLQDIITAILMDGKDDLSDDKLTVARAKIQRFLSQPFVVAEILFTGAPGKI
	/number=5		complement(19850..19963)
	17399..17477	exon	/gene="At5g08670"
	/gene="At5g08660"		/number=1
exon	/number=5	intron	complement(19964..20203)
	17478..17575		/gene="At5g08670"
	/gene="At5g08660"	exon	/number=1
intron	/number=6		complement(20204..20515)
	17576..17766		/gene="At5g08670"
	/gene="At5g08660"	intron	/number=2
exon	/number=6		complement(20516..20609)
	17767..17836		/gene="At5g08670"
	/gene="At5g08660"	exon	/number=2
intron	/number=7		complement(20610..20834)
	17837..18027		/gene="At5g08670"
	/gene="At5g08660"	intron	/number=3
exon	/number=7		complement(20835..20917)
	18028..18131		/gene="At5g08670"
	/gene="At5g08660"	intron	/number=3
intron	/number=8		
	18132..18230		



```

exon      complement (20918. .21076)
/gene="At5g08670"
number=4
intron    complement (21077. .21176)
/gene="At5g08670"
number=4
exon      complement (21177. .21280)
/gene="At5g08670"
number=5
intron    complement (21281. .21359)
/gene="At5g08670"
number=5
exon      complement (21360. .21477)
/gene="At5g08670"
number=6
intron    complement (21478. .21568)
/gene="At5g08670"
number=6
exon      complement (21569. .21657)
/gene="At5g08670"

```

## Alignment Scores:

```

Pred. No.:      1.57e-05      Length:      82896
Score:          174.00      Matches:      31
Percent Similarity: 90.24%      Conservative: 6
Best Local Similarity: 75.61%      Mismatches: 4
Query Match:      32.83%      Indels: 0
DB:              8      Gaps: 0

```

US-10-629-953-4 (1-101) x ATT2K12 (1-82896)

```

Qy      35 ILeValValAlaGlnIleSerLysAsnArgValAlaValAlaArgThrTrpAsnGlyLys 54
Db      77551 TTGGTTATACATCATGATATCTAGAAATAGGAGAGTCTCTGTAGCAATTCGAACCGGAAG 77492

Qy      55 ValValValAspIleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLys 74
Db      77491 ATTGGATTGCATTCGTGAGTCTCTATGTCAGGACCGAAAGACTTTTGCTGGCAAGAA 77432

Qy      75 Gly 75
Db      77431 GGT 77429

```

Search completed: January 12, 2005, 02:38:40  
 Job time : 3329 secs

**THIS PAGE BLANK (USP10)**

**THIS PAGE BLANK (USP10)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 11, 2005, 23:18:54 ; Search time 412 Seconds

(without alignments)  
1286.873 Million cell updates/sec

Title: US-10-629-953-4

Perfect score: 530

Sequence: 1 MWKGNRFGGGEPAKRR.....QWKLRDNKAIDEAIKENA 101

Scoring table:

	BLOSUM62	Xgapop 10.0	Xgapext 0.5
Xgapop 10.0			
Xgapop 10.0			
Xgapop 6.0			
Delop 6.0			

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/h/US10629953/runat 10012005 172534 9407/app_query.fasta_1.263
-DB=N_Geneseq 23Sep04 -QPM=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOPCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=spct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10629953 @CGN 1 1 470 @runat 10012005 172534 9407 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

	N Geneseq 23Sep04:*
1:	Geneseq1980s:*
2:	Geneseq1990s:*
3:	Geneseq2000s:*
4:	Geneseq2001as:*
5:	Geneseq2001bs:*
6:	Geneseq2002as:*
7:	Geneseq2002bs:*
8:	Geneseq2003as:*
9:	Geneseq2003bs:*
10:	Geneseq2003cs:*
11:	Geneseq2003ds:*
12:	Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Length	ID	Description
1	530	100.0	632 3	AAZ50412 Rice tran
2	398	75.1	649 3	AAZ50411 Corn tran
3	389.5	73.5	460 3	AAZ50414 Wheat tra
4	368	69.4	289 10	ABX82999 Corn ear-
5	304.5	57.5	310 6	ABL70693 Corn tass
6	297	56.0	263 10	ABX83152 Corn ear-

7	275	51.9	286	10	ABX88456	Abx88456 Corn ear-
8	269	50.8	672	3	AAZ50413	Aaz50413 Soybean t
9	247.5	46.7	471	3	AAZ38126	Aaz38126 Arabidops
10	206.5	39.0	516	3	AAZ50416	Aaz50416 Vernonia
11	194.5	36.7	1063	3	AAZ32760	Aac32760 Arabidops
12	185	34.9	913	8	ACF03526	Acf03526 Brassica
13	184.5	34.8	740	3	AAZ50419	Aaz50419 Soybean t
c	183	34.5	496	10	ADB81749	Ad81749 Arabidops
14	183	34.5	757	3	AAZ50198	Aaz50198 Arabidops
15	183	34.5	757	3	AAZ34846	Aac34846 Arabidops
16	183	34.5	1089	3	AAZ50417	Aaz50417 Corn tran
17	168.5	31.8	498	3	AAZ50415	Aaz50415 Marigold
18	167.5	31.6	615	10	ADB57912	Ad57912 Toxicity-
19	166	31.3	615	10	ADB57912	Ad57912 Toxicity-
20	166	31.3	939	3	AAZ50418	Aaz50418 Corn tran
21	166	31.3	939	3	AAZ50418	Aaz50418 Corn tran
22	163	30.8	483	9	ACH38105	Ach38105 Human end
23	160.5	30.3	384	5	ABA83091	Ab83091 Polymetas
24	160.5	30.3	384	8	ACC50254	Aac50254 Breast ca
25	160.5	30.3	384	10	ADF76607	Adf76607 Novel hum
26	160.5	30.3	443	8	ABZ56521	Abz56521 Aspergill
27	160.5	30.3	460	9	ACH47965	Ach47965 Human lun
28	160.5	30.3	592	6	ABK16067	Abk16067 Human lun
29	160.5	30.3	592	10	ADB95330	Ad95330 Human lun
30	160.5	30.3	601	6	ABK16230	Abk16230 Human lun
31	160.5	30.3	601	10	ADB95493	Ad95493 Human lun
32	160.5	30.3	674	4	AAD23497	Aad23497 Human lun
33	160.5	30.3	674	10	ADD66785	Add66785 Human lun
34	160.5	30.3	674	10	ADB88039	Ad88039 Human lun
35	160.5	30.3	689	6	ABL42297	Ab142297 Pancreati
36	160.5	30.3	701	2	AAZ77517	Aaz77517 Human ova
37	160.5	30.3	1336	12	ADU91493	Ad191493 Human imm
38	160.5	30.3	1336	12	ADN04675	Adn04675 Antipeori
39	160.5	30.3	1391	5	ABV25741	Abv25741 Human pro
40	160.5	30.3	3527	3	AAZ99092	Aac99092 Human pan
c	160	30.2	622	6	ABK83518	Abk83518 Human CDN
42	157.5	29.7	432	9	ACH19785	Ach19785 Human adu
43	154.5	29.2	493	5	AAZ91456	Aaz91456 DNA encod
44	153.5	29.0	662	3	AAZ10376	Aac10376 Human sec
c	147.5	27.8	547	12	ADP03062	Adp03062 Human hou

#### ALIGNMENTS

RESULT 1

AAZ50412  
ID AAZ50412 standard; cDNA; 632 BP.

XX  
AC AAZ50412;

XX  
DT 18-MAY-2000 (first entry)

XX  
DE Rice transcription coactivator PC4 (P15) type 1 cDNA.

XX  
KW Positive cofactor 4; PC4; transcription coactivator; PC4 (P15) type 1;  
KW rice; signal mediator; activator; Gal4/VPI6; transgenic plant;  
KW general transcription factor; GTF; transcription initiation complex;  
KW immunological screening; detection; marker; clone rrl.pk0003.a12; ss.

XX  
OS Oryza sativa.

XX  
FH Key Location/Qualifiers

FT CDS 46..351

FT FT /\*tag= a

FT FT /product= "Rice PC4 (P15) type 1 transcription

FT FT coactivator"

FT FT /note= "Derived from clone rrl.pk0003.a12"

XX  
PN WO200005377-A2.

XX  
XX 03-FEB-2000.

XX  
XX 21-JUL-1999; 99WO-US016479.

XX  
PF

```
PR 22-JUL-1998; 98US-0093687P.
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
PA Cahoon RE, Caiami PG, Odell JT, Sakai H, Zhu Q;
XX WPI; 2000-182701/16.
DR P-PSDB; AAY44879.
XX
XX Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT used to alter the level of PC4(P15) type 1 and PC4(P15) type 2
PT polynucleotides and polypeptides.
XX
XX Claim 9; Page 33; 43pp; English.
XX
XX The present sequence is the cDNA encoding rice Positive Cofactor 4 (PC4)
CC transcription coactivator, designated as PC4(P15) type 1. It is isolated
CC from clone rri.pk0003.al2, obtained from rri cDNA library prepared using
CC rice root two week old developing seedlings. PC4 molecules functions as a
CC signal mediator between activators like Gal4/VP16 and general
CC transcription factors (GTFs) in a transcription initiation complex. The
CC expression levels of PC4 can be manipulated and the functional properties
CC of specific transcriptional activators can be modulated. The PC4 DNA can
CC be used to create transgenic plants with altered PC4 levels, that would
CC affect the level of transcription of specific genes in the plant. It is
CC also used for immunological screening of cDNA libraries and to raise
CC specific antibodies for detection. The DNA is used as probes and primers,
CC for genetic and physical mapping of genes and as markers for traits
CC linked to those genes
XX
SQ Sequence 632 BP; 165 A; 139 C; 187 G; 141 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6,69e-54 Length: 632
Score: 530.00 Matches: 101
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-629-953-4 (1-101) x AAZ50412 (1-632)
QY 1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGlyGluProAlaAlaLysArgArg 20
DB 46 ATGTGGCGGAAGGGAACAAGCGTTTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 105
QY 21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspLeValValAlaGlnIle 40
DB 106 GCGCGCGCGGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGATA 165
QY 41 SerLysAsnArgArgValAlaValAtgThrTrpAsnGlyLysValValValAspIleArg 60
DB 166 TCGAAGACAGAGAGGTTGGCGGTTCGGACCTCGACGCGCGCGCGCGCGCGCGCGCG 225
QY 61 GluPheTrpGlyLysAspGlyLysThrLeuProGlyArgLysGlyLysGlnLeuProMet 80
DB 226 GAGTTCACGAGAGGAGCGCGCAAGACCTCCCGCGCGCGCAAGGTATACAGCTCCCAATG 285
QY 81 AspGlnTrpLysLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100
DB 286 GATCAGTGAAGACTACTGAGGAGCAATATCAAGCTATAGATGAGGCGCATCAAGGAGAA 345
QY 101 Ala 101
DB 346 GCG 348

RESULT 2
ID AAZ50411
XX AAZ50411 standard; cDNA; 649 BP.
AC AAZ50411;
XX
DT 18-MAY-2000 (first entry)

Corn transcription coactivator PC4(P15) type 1 cDNA.
Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 1;
corn; signal mediator; activator; Gal4/VP16; transgenic plant;
general transcription factor; GTF; transcription initiation complex;
immunological screening; detection; marker; clone cca.pk0020.d2; ss.
Zea mays.
XX
XX Key Location/Qualifiers
FH CDS 33..344
FT /*tag= a
FT /product= "Corn PC4(P15) type 1 transcription
FT coactivator"
FT /note= "Derived from clone cca.pk0020.d2"
XX
XX WO200005377-A2.
XX
XX 03-FEB-2000.
XX
XX 21-JUL-1999; 99WO-US016479.
XX
XX 22-JUL-1998; 98US-0093687P.
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX Cahoon RE, Caiami PG, Odell JT, Sakai H, Zhu Q;
XX
XX WPI; 2000-182701/16.
DR P-PSDB; AAY44878.
XX
XX Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT used to alter the level of PC4(P15) type 1 and PC4(P15) type 2
PT polynucleotides and polypeptides.
XX
XX Claim 3; Page 32; 43pp; English.
XX
XX The present sequence is the cDNA encoding corn Positive Cofactor 4 (PC4)
CC transcription coactivator, designated as PC4(P15) type 1. It is isolated
CC from the corn callus type II tissue (undifferentiated) from the cDNA
CC clone cca.pk0020.d2. PC4 molecules functions as a signal mediator between
CC activators like Gal4/VP16 and general transcription factors (GTFs) in a
CC transcription initiation complex. The expression levels of PC4 can be
CC manipulated and the functional properties of specific transcriptional
CC activators can be modulated. The PC4 DNA can be used to create transgenic
CC plants with altered PC4 levels, that would affect the level of
CC transcription of specific genes in the plant. It is also used for
CC immunological screening of cDNA libraries and to raise specific
CC antibodies for detection. The DNA is used as probes and primers, for
CC genetic and physical mapping of genes and as markers for traits linked to
CC those genes
XX
SQ Sequence 649 BP; 167 A; 147 C; 182 G; 153 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4,04e-38 Length: 649
Score: 398.00 Matches: 76
Percent Similarity: 87.25% Conservative: 13
Best Local Similarity: 74.51% Mismatches: 11
Query Match: 75.09% Indels: 2
DB: 3 Gaps: 1

US-10-629-953-4 (1-101) x AAZ50411 (1-649)
QY 1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGly-----GlyGluProAlaAlaLys 18
DB 33 ATGTGGGGAAGGAAGAAAGCGTTTCGGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCAAG 92
QY 19 ArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspIleValValAla 38
DB 93 CGCCAGGCGCGGAGGAGCGATGCGCCCTCCGAATCCGCTGAAGACGGTACCGTCGTAGCC 152
```

QY 39 GlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAsp 58  
 Db 153 GAGATATCGAAGAACAAAGAGGTGTCGTTAGGAGCTGGAAGCGCAGGCTTCGTGTCGAC 212  
 QY 59 IleArgGluPheTyTGlyLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeu 78  
 Db 213 TTACGGAGTTCTACTTCAGGAGCGGAGACTCTCCCAACCGCAAGGTATATCACTC 272  
 QY 79 ProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLys 98  
 Db 273 CAATTAGATCAGTGTGAAGTACTGAAGGACCAACATCAAGCCATAAATGAGGCAATCGAG 332  
 QY 99 GluAsn 100  
 Db 333 GAAAT 338  
 RESULT 3  
 ID AAZ50414 standard; cDNA; 460 BP.  
 XX AAZ50414;  
 AC AAZ50414;  
 XX 18-MAY-2000 (first entry)  
 DT Wheat transcription coactivator PC4 (P15) type 1 cDNA.  
 DE  
 XX Positive cofactor 4; PC4; transcription coactivator; PC4 (P15) type 1;  
 KW wheat; signal mediator; activator; Gal4/Vp16; transgenic plant;  
 KW general transcription factor; GTF; transcription initiation complex;  
 KW immunological screening; detection; marker; clone wdk2c.pk015.g20; ss.  
 XX  
 OS Triticum aestivum.  
 XX  
 XX Key Location/Qualifiers  
 FH 3. .293  
 CD  
 FT /\*tag= a  
 FT /product= "Wheat PC4 (P15) type 1 transcription  
 FT coactivator"  
 FT /note= "Derived from clone wdk2c.pk015.g20"  
 XX  
 XX W0200005377-A2.  
 XX  
 XX 03-FEB-2000.  
 XX  
 XX 21-JUL-1999; 99WO-US016479.  
 XX  
 XX 22-JUL-1998; 98US-0093687P.  
 XX  
 XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
 XX  
 XX Cahoon RE, Caimi PG, Odell JT, Sakai H, Zhu Q;  
 XX WPI; 2000-182701/16.  
 XX P-PSDB; AAY44881.  
 XX  
 XX Novel PC4 transcriptional coactivator polynucleotides and polypeptides  
 PT used to alter the level of PC4 (P15) type 1 and PC4 (P15) type 2  
 PT polynucleotides and polypeptides.  
 XX  
 XX Claim 3; Page 35; 43pp; English.  
 XX  
 CC The present sequence is the cDNA encoding wheat Positive Cofactor 4 (PC4)  
 CC transcription coactivator, designated as PC4 (P15) type 1. It is isolated  
 CC from clone wdk2c.pk015.g20, obtained from wdk2c cDNA library, prepared  
 CC from wheat developing kernels, 7 days after anthesis. PC4 molecules  
 CC functions as a signal mediator between activators like Gal4/Vp16 and  
 CC general transcription factors (GTFs) in a transcription initiation  
 CC complex. The expression levels of PC4 can be manipulated and the  
 CC functional properties of specific transcriptional activators can be  
 CC modulated. The PC4 DNA can be used to create transgenic plants with  
 CC altered PC4 levels, that would affect the level of transcription of  
 CC specific genes in the plant. It is also used for immunological screening  
 CC of cDNA libraries and to raise specific antibodies for detection. The DNA

CC is used as probes and primers, for genetic and physical mapping of genes  
 CC and as markers for traits linked to those genes  
 XX  
 SQ Sequence 460 BP; 117 A; 110 C; 139 G; 87 T; 0 U; 7 Other;  
 Alignment Scores: 2.72e-37 Length: 460  
 Pred. No.: 389.50 Matches: 76  
 Score: 87.37% Conservative: 7  
 Percent Similarity: 80.00% Mismatches: 9  
 Best Local Similarity: 73.49% Indels: 3  
 Query Match: 3  
 DB: 1  
 US-10-629-953-4 (1-101) x AAZ50414 (1-460)  
 QY 9 PheGlyGlyGly-----GlyGluProAlaAlaLysArgAlaAlaGlyAspAsp 25  
 Db 3 TTCGGCGGCG 62  
 QY 26 GlyProSerGluSerAlaAspAspIleValAlaGlnIleSerLysAsnArgArg 45  
 Db 63 GGTCCCTCCGAGGAAACCGACGCGCATCGTTCGTCGATATCGAAGAACAGAGG 122  
 QY 46 ValAlaValArgThrTrpAsnGlyLysValValValValValValValValValVal 65  
 Db 123 GTGCGCGTGAGGAACCTGGAACCGGAAGGTCATCGTCGACATCGCGGAGTTCTACGAAAG 182  
 QY 66 AspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMetAspGlnTrpLysIle 85  
 Db 183 GACGGCAAGAGCTCCGACCCCGCAAGGTATATCGTCTCAATGATCATGTGGAATA 242  
 QY 86 LeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100  
 Db 243 CTCGAGGCAACATCATCGAAGCTATAGACGAGGCCATCAAGGAGAAC 287  
 RESULT 4  
 ABX82999  
 ID ABX82999 standard; cDNA; 289 BP.  
 XX  
 AC ABX82999;  
 XX  
 DT 24-APR-2003 (first entry)  
 XX  
 DE Corn ear-derived polynucleotide (cpd) #1459.  
 XX  
 KW Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; SATMON023;  
 KW structural gene; functional gene; regulatory gene;  
 KW corn ear-specific profile; gene transcription; gene expression;  
 KW hybrid plant; desirable trait expression; plant breeding program;  
 KW inheritance; desirable characteristic; growth; development;  
 KW disease resistance; environmental adaptability; quality; yield;  
 KW multigene trait; plant; gene; ss.  
 XX  
 OS Zea mays.  
 XX  
 XX US6476212-B1.  
 XX  
 XX 05-NOV-2002.  
 XX  
 XX 14-MAY-1999; 99US-00313294.  
 XX  
 XX 26-MAY-1998; 98US-0086722P.  
 XX  
 XX (INCY-) INCYTE GENOMICS INC.  
 XX  
 XX Lalgudi RV, Ito LY, Sherman BK;  
 XX WPI; 2003-208840/20.  
 XX  
 XX Novel purified corn-ear derived polynucleotide useful as hybridization  
 XX probe for detecting polynucleotide in sample, and for identifying,  
 XX evaluating, and altering desired characteristics associated with growth,  
 XX development.

XX Example; SEQ ID NO 1459; 390pp; English.

XX The present invention relates to the isolation of corn ear-derived

XX polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022

XX and SATMON023. Some of the cdfs uniquely identify structural, functional,

XX and regulatory genes of corn ear. The polynucleotide sequences are

XX useful for detecting cdfs in a sample, for producing a corn ear-specific

XX profile of gene transcription, for detecting altered gene expression in

XX inbred or hybrid plants, and for screening several molecules for specific

XX binding to the polynucleotide. The cdfs are useful to identify, isolate,

XX or extend identical or related corn-ear nucleic acid sequences from DNA

XX libraries, and in nucleic acid amplification or hybridisation techniques

XX to follow the expression of desirable traits through plant breeding

XX programs. Preferably, the cdfs are used to identify, evaluate, alter, or

XX follow the inheritance of desired characteristics associated with growth

XX and development, disease resistance, environmental adaptability, quality,

XX and yield of corn. The cdfs are also useful as molecular markers for

XX studying inheritance and multigene traits in a plant breeding program.

XX The cdfs are useful for producing purified corn-ear polypeptides by

XX recombinant techniques. They are also useful in diagnostic assays to

XX detect or confirm conditions or diseases associated with abnormal levels

XX of cdp expression. ABX81541-ABX9140 represent corn ear-derived

XX polynucleotides (cdps) of the invention. Note: The sequence data for this

XX patent did not form part of the printed specification, but was obtained

XX in electronic format directly from the USPTO web site at

XX seqdata.uspto.gov/psipdIDEntry.html

XX SQ Sequence 289 BP; 79 A; 71 C; 92 G; 47 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5,63e-35	Length:	289
Score:	368.00	Matches:	70
Percent Similarity:	86.17%	Conservative:	11
Best Local Similarity:	74.47%	Mismatches:	11
Query Match:	69.43%	Indels:	2
DB:	10	Gaps:	1

US-10-629-953-4 (1-101) x ABX82999 (1-289)

QY 1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGly-----GlyGluProAlaAlaLys 18

DB 6 ATGTGGGGGAAAGAAAGAGCGTTTCGGCGGTGGCGGAGCGGTGAGCCGCGGCCAAG 65

QY 19 ArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspIleValValAla 38

DB 66 CGCAGCCCGGAGGAGGACGATGCGCCCTCCGAATCGCTGAAGACGGTACCGTCGTAGCC 125

QY 39 GlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValAsp 58

DB 126 GAGATATCGAAGAACAGAGAGGTTCGTTAGGAGCTGGAAGACGAGGTCTTCGTGCAC 185

QY 59 IleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeu 78

DB 186 TTACGCGAGTCTTACTTCAAGGACGCGACAGACTCTCCACCCGCAAGGTATATCACTC 245

QY 79 ProMetAspGlnTrpIlyIleLeuArgAspAsnIleLysAla 92

DB 246 CAATTAGATCAGTGGGAAGATATTGAAGGACAACATCAAGCC 287

RESULT 5

ID ABL70693 standard; CDNA; 310 BP.

XX ABL70693;

XX 14-MAY-2002 (first entry)

XX Corn tassel-derived polynucleotide (cdps) SEQ ID NO:67.

XX Corn; corn tassel-derived polynucleotide; cdfs; hybrid breeding; CDPs;

XX inheritance; characteristic; growth; development; disease resistance;

XX environmental adaptability; quality; yield; yield; molecular marker;

KW multigene trait; plant breeding; corn tassel; gene; ss.

OS Zea mays.

XX US2001051335-A1.

PN 13-DEC-2001.

PD 16-APR-1999; 99US-00294093.

PF 21-APR-1998; 98US-0082567P.

XX (LALG/) LALGUDI R V.

PA (ITOL/) ITO L Y.

PA (SHER/) SHERMAN B K.

XX Lalgudi RV, Ito LY, Sherman BK;

PI WPI; 2002-163647/21.

DR Novel purified corn tassel-derived polynucleotide useful for determining

PT altered gene expression, to recover regulatory elements and to follow

PT inheritance of desirable characteristics through hybrid breeding

PT programs.

XX Claim 1; SEQ ID NO 67; 201pp; English.

XX The present sequence describes a purified corn tassel-derived

CC polynucleotide sequence (cdps) comprising a nucleic acid sequence

CC selected from those given in ABL70627 to ABL76833. The cdfs sequences

CC encode corn tassel-derived polypeptides (CDPs). The cdfs sequences (I)

CC can be used for determining altered gene expression, to recover

CC regulatory elements and to follow inheritance of desirable

CC characteristics through hybrid breeding programs. (I) are also useful in

CC the evaluation, and alteration of desired characteristics associated with

CC growth and development, disease resistance, environmental adaptability,

CC quality and yield, and as molecular markers for studying inheritance of

CC multigene traits in a plant breeding program. (I) can be used to produce

CC a tassel-specific profile of gene transcription, a transcript image, to

CC clone regulatory elements for use in transformation vectors, to express a

CC polypeptide, to identify, isolate or extend identical or related corn

CC tassel nucleic acid sequences from DNA libraries, in nucleic acid

CC hybridisation or amplification technologies, as query sequences to

CC determine homology of known sequences, as probe for use in Southern or

CC Northern hybridisation, and to identify the presence of and/or to

CC determine the degree of similarity between two (or more) nucleic acid

CC sequences

XX SQ Sequence 310 BP; 84 A; 77 C; 96 G; 50 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.:	2,36e-27	Length:	310
Score:	304.50	Matches:	65
Percent Similarity:	81.44%	Conservative:	14
Best Local Similarity:	67.01%	Mismatches:	13
Query Match:	57.45%	Indels:	5
DB:	6	Gaps:	2

US-10-629-953-4 (1-101) x ABL70693 (1-310)

QY 1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGly-----GlyGluProAlaAlaLys 18

DB 26 ATGTGGGGGAGGAAAGAAAGCGTTTCGGCGGTGGCGGAGCGCGGCGGCCAAG 85

QY 19 ArgArgAlaAla-GlyAspAspGlyProSerGluSerAlaAspAspIleValValAl 38

DB 86 CGCAGCCCGGAGGAGGACGATGCGCCCTCCGAATCGCTGAAGACGGTACCGTCGTAGC 145

QY 38 aGlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValAs 58

DB 146 CGAGATATCGAAGAACAGAGGTATCCGTTAGGAGCTGGAAGGAGCGGTCTTCGTGCA 205

QY 58 pIleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGln-L 78







```
Db          |||||:::
          343 GAGAAATTCT 351
RESULT 9
AAC38126
ID AAC38126 standard; DNA; 471 BP.
XX
AC AAC38126;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 19870.
XX
KW Hybridisation assay; Genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN BP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126284P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138554P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147433P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
```





```
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145919P.
PR 02-AUG-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 03-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 08-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.

PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160816P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 1.52e-13
Score: 194.50
Percent Similarity: 70.49%
Best Local Similarity: 63.93%
Query Match: 36.70%
DB: 3
Gaps: 1

US-10-629-953-4 (1-101) x AAC32760 (1-1063)
QY 15 ProAlaAlaLysArgAlaAlaGlyAspGlyProSerGluSerAlaAspAsp 34
DB 175 CCGGCTAAGAAAGTCGAAAGCCAGCAGCAGAT-----TCCGACCAGTCTGACGAT 225
QY 35 IleValValAlaGlnIleSerLysAenArgArgValAlaValArgThrTrpAsnGlyLys 54
DB 226 ATCGTCGTCTGCAACATATCTAAGAAATAGGAGAGTCTCTGTAAAGGAATTGGAACGGGAAG 285
QY 55 ValValValAspIleArgGluPheTyGluLysAspGlyLysThrLeuProGlyArgLys 74
DB 286 ATTTGGATTGACATTCGTGAGTTCTATGTCNAGGACGGAAGACTTTGCTCTGCGCAGAA 345
QY 75 Gly 75
DB 346 GGT 348

RESULT 12
ACF03526
ID ACF03526 standard; cDNA; 913 BP.
XX AC ACF03526;
XX AC ACF03526;
DT 27-OCT-2003 (revised)
DT 12-SEP-2003 (first entry)
XX DE Brassica campestris MIP102 encoding cDNA SEQ ID NO:1.
XX KW Brassica campestris; plant; MIP102; virus resistance;
XX plant virus transport protein binding protein; gene; ss.
XX OS Brassica rapa.
XX FH Key Location/Qualifiers
XX CDS 14..511
XX /tag= a
XX /product= "MIP102"
XX WO2003022039-A1.
XX PN
XX PD 20-MAR-2003.
XX PF 10-SEP-2001; 2001WO-JP007858.
XX XX
```



Db 376 GCGAAGAACTTCTACTTCCAAAGGAATTAAGTTTGACAGAGAGCAGTGGTCAGCGCTTT 435  
 Qy 87 ArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99  
 Db 436 AAGAAAATGTGCTGCGCATAGAAAAGCCATTAAAGAAA 474  
 RESULT 14  
 ADEB1749/c  
 ID ADEB1749 standard; cDNA; 496 BP.  
 AC ADEB1749;  
 XX  
 XX 29-JAN-2004 (first entry)  
 DT  
 DE Arabidopsis thaliana expressed polynucleotide seq id 520.  
 XX  
 XX genetically modified organism; transgenic organism; plant;  
 KW inhibitor testing; activator testing; modifier testing; fungicide;  
 KW insecticide; genetic function; genetic regulation; cellular metabolism;  
 KW gene; ss.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 XX US2003115639-A1.  
 XX  
 XX 19-JUN-2003.  
 XX  
 XX 26-JAN-2001; 2001US-00770961.  
 XX  
 XX 27-JAN-2000; 2000US-0178466P.  
 XX  
 XX (GORL/) GORLACH J.  
 PA (ANY/) AN Y.  
 PA (HAMI/) HAMILTON C M.  
 PA (PRIC/) PRICE J L.  
 PA (RAIN/) RAINES T M.  
 PA (YUY/) YU Y.  
 PA (RAME/) RAMEAKA J G.  
 PA (PAGE/) PAGE A.  
 PA (MATH/) MATHAW A V.  
 PA (LEDF/) LEDFORD B L.  
 PA (WOES/) WOESSNER J P.  
 PA (HAAS/) HAAS W D.  
 PA (GARC/) GARCIA C A.  
 PA (KRIC/) KRICKER M.  
 PA (SLAT/) SLATER T.  
 PA (DAVI/) DAVIS K R.  
 PA (ALLE/) ALLEN K.  
 PA (HOFF/) HOFFMAN N.  
 PA (HURB/) HURBAN P.  
 XX  
 XX Gorchach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y,  
 PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;  
 PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;  
 PI Hurban P;  
 XX  
 XX WPI; 2003-810930/76.  
 XX  
 XX Novel Arabidopsis thaliana nucleic acids useful for generating  
 PT genetically modified transgenic organisms, for screening biologically  
 PT active agents such as fungicides, insecticides.  
 XX  
 XX Claim 1; SEQ ID NO 520; 44pp; English.  
 XX  
 XX The invention describes a nucleic acid (I) comprising a sequence capable  
 CC of hybridising under stringent conditions to any one of 999 fully defined  
 CC Arabidopsis thaliana sequences (I) as given in specification e.g., 360,  
 CC 1137, 455, 219, 472 nucleotides, etc, or its fragment. (I) is useful as a  
 CC hybridisation probe to complementary molecules in a cDNA library. (I) is  
 CC also useful for generating genetically modified and transgenic organisms,  
 CC usually plant cells and plants. A protein encoded by (I) is useful in  
 CC screening assays to determine the effect of candidate inhibitors,  
 CC activators or modifiers of the gene product. The protein is also useful

CC for screening biologically active agents e.g., fungicides and  
 CC insecticides. A genetically modified cell, comprising an exogenous  
 CC nucleic acid, where the nucleic acid comprises transcription regulatory  
 CC sequences operably linked to a sequence capable of hybridising under  
 CC stringent conditions to (I) is useful in the study of genetic function  
 CC and regulation, for alteration of the cellular metabolism and for  
 CC screening compounds that may affect the biological function of the gene  
 CC or gene product. This sequence represents an Arabidopsis thaliana  
 CC polynucleotide of the invention.  
 XX  
 SQ Sequence 496 BP; 147 A; 108 C; 77 G; 164 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 1.38e-12 Length: 496  
 Score: 183.00 Matches: 37  
 Percent Similarity: 56.25% Conservative: 17  
 Best Local Similarity: 38.54% Mismatches: 22  
 Query Match: 34.53% Indels: 20  
 DB: 10 Gaps: 2  
 US-10-629-953-4 (1-101) x ADEB1749 (1-496)  
 Qy 4 LysGlyAsnLysArgPheGlyGlyGlyGluProAlaAlaLysArgAlaAlaGly 23  
 Db 459 AAAGAAACAAAGAGTTT----- 442  
 Qy 24 AspAspGlyProSerGluSerAlaAspAspAspIleValAlaGlnIleSerLysAsn 43  
 Db 441 GATGAT-----GACGGCATCTTATCATTTGCAGTTATCGGATAAG 400  
 Qy 44 ArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArgGluPheTyr 63  
 Db 399 AGAAGAGTGACGATTTCAGGAATTTAAAGGAGAGAGTTTGGTTCTTATCAGAGAGTATTAC 340  
 Qy 64 GluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMetAspGlnTyr 83  
 Db 339 AAGAAAGATGGCAAGAACTTCTACTTCTTAAAGGAATAAGCTTAAACAGATGAACAATGG 280  
 Qy 84 LysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99  
 Db 279 TCAACCTTCAAGAAACATATGCAGCCATCGAAATGCTGTCAAGAAA 232  
 RESULT 15  
 AAC50198  
 ID AAC50198 standard; DNA; 757 BP.  
 XX  
 AC AAC50198;  
 XX  
 XX 18-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 63940.  
 XX  
 XX Hybridisation assay; genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway; metabolic pathway;  
 KW promoter; termination sequence; ss.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 XX EF1033405-A2.  
 PN  
 XX  
 PD 06-SEP-2000.  
 XX  
 XX 25-FEB-2000; 2000EP-00301439.  
 PF  
 XX 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.



PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

## Alignment Scores:

Pred. No.:	2.34e-12	Length:	757
Score:	183.00	Matches:	37
Percent Similarity:	56.25%	Conservative:	17
Best Local Similarity:	38.54%	Mismatches:	22
Query Match:	34.53%	Indels:	20
DB:	3	Gaps:	2

US-10-629-953-4 (1-101) x AAC50198 (1-757)

Qy	4	LysGlyAsnLysArgPheGlyGlyGlyGluProAlaAlaLysArgAlaAlaGly	23
Db	343	AAAGGAAACAAAGAGTTT-----	360
Qy	24	AspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLysAsn	43
Db	361	GATGAT-----GACGGCGATCTTATCATCTTTCAGGTTATCGGATAAG	402
Qy	44	ArgArgValAlaValArgThrTyrAsnGlyLysValValValAspIleArgGluPheTyr	63
Db	403	AGAAAGTGACGATTCAGGAATTTAAAGGGAAGAGTTTGTTCTATCAGAGAGTATTAC	462
Qy	64	GluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMetAspGlnTyr	83
Db	463	AAGAAAGATGCAAGAACTTCCTACTTCTAAAGGATAGCTTAACAGATGAACATGG	522
Qy	84	LysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu	99
Db	523	TCAACCTTCAAGAAACATGCCACCATCGAAAATGCTGTCAAGAAA	570

Search completed: January 12, 2005, 02:45:17  
Job time : 417 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2005, 01:26:01 ; Search time 88 Seconds  
(without alignments)  
815.792 Million cell updates/sec

Title: us-10-629-953-4  
Perfect score: 530  
Sequence: 1 MWKGNKRFQGGGEPAAKRR.....QWKILRDNKAIDEAIKENA 101

Scoring table: BLOSUM62 Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool\_h/US10629953/runat\_10012005.172536.9437/app.query.fasta\_1.263  
-DB=issued\_Patents\_NA -QMT=fastcap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10629953@cgn 1.1.69 @runat\_10012005.172536.9437 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=10 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	368	69.4	289	US-09-313-294A-1459	Sequence 1459, Ap
2	297	56.0	263	US-09-313-294A-1612	Sequence 1612, Ap
3	275	51.9	286	US-09-313-294A-6916	Sequence 6916, Ap
4	160.5	30.3	674	US-09-854-133-477	Sequence 477, App
5	153.5	29.0	662	US-09-513-999C-14451	Sequence 14451, A
6	136.5	25.8	452	US-09-621-976-18629	Sequence 18629, A
7	106	20.0	480	US-09-248-796A-5030	Sequence 5030, Ap
8	84	15.8	325	US-09-621-976-18626	Sequence 18626, A
9	84	15.8	400	US-09-513-999C-14452	Sequence 14452, A
10	84	15.8	479	US-09-621-976-18624	Sequence 18624, A
11	79.5	15.0	1320	US-09-724-797-85	Sequence 85, Appl
12	75	14.2	2058	US-09-252-991A-15140	Sequence 15140, A

13	75	14.2	2211	4	US-09-252-991A-14623	Sequence 14623, A
c 14	75	14.2	2712	4	US-09-919-172-40	Sequence 40, Appl
c 15	74.5	14.1	528	4	US-09-489-039A-1584	Sequence 1584, Ap
c 16	74.5	14.1	1230	4	US-09-489-039A-1484	Sequence 1484, Ap
c 17	74.5	14.1	1737	4	US-09-489-039A-1730	Sequence 1730, Ap
c 18	74	14.0	828	4	US-09-543-681A-3276	Sequence 3276, Ap
c 19	74	14.0	1524	4	US-09-489-039A-5551	Sequence 5551, Ap
c 20	73.5	13.9	1224	4	US-09-266-965-22	Sequence 22, Appl
c 21	73.5	13.9	12249	4	US-09-266-965-74	Sequence 74, Appl
c 22	73.5	13.9	18331	4	US-09-266-965-96	Sequence 96, Appl
c 23	73.5	13.9	640681	4	US-09-790-988-1	Sequence 1, Appli
c 24	73.5	13.9	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c 25	73.5	13.9	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c 26	72.5	13.7	4043	4	US-09-814-915A-70	Sequence 13408, A
c 27	72	13.6	226	4	US-09-621-976-13408	Sequence 147, App
c 28	72	13.6	547	4	US-09-513-999C-148	Sequence 147, App
c 29	72	13.6	649	4	US-09-513-999C-147	Sequence 147, App
c 30	72	13.6	714	4	US-09-252-991A-674	Sequence 674, App
c 31	72	13.6	951	4	US-09-252-991A-741	Sequence 741, App
c 32	71.5	13.5	16826	4	US-08-956-171E-46	Sequence 46, Appl
c 33	71.5	13.5	16826	4	US-08-781-986A-46	Sequence 15005, A
c 34	71	13.4	474	4	US-09-252-991A-15005	Sequence 15872, A
c 35	70.5	13.3	837	4	US-09-252-991A-15872	Sequence 1330, Ap
c 36	70.5	13.3	972	4	US-09-248-796A-1330	Sequence 15820, A
c 37	70.5	13.3	1158	4	US-09-252-991A-15820	Sequence 351, App
c 38	70.5	13.3	2366	4	US-09-799-451-351	Sequence 15898, A
c 39	70.5	13.3	2970	4	US-09-252-991A-15898	Sequence 2, Appli
c 40	70	13.2	5173	3	US-08-801-308-2	Sequence 2, Appli
c 41	70	13.2	4403765	3	US-09-103-840A-2	Sequence 1, Appli
c 42	70	13.2	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c 43	69.5	13.1	162450	3	US-09-345-882-1	Sequence 9, Appli
c 44	69	13.0	10709	4	US-09-596-002-9	Sequence 1, Appli
c 45	68.5	12.9	43280	2	US-08-804-227C-1	

ALIGNMENTS

RESULT 1  
US-09-313-294A-1459  
; Sequence 1459, Application US/09313294A  
; Patent No. 6476212  
; GENERAL INFORMATION:  
; APPLICANT: Lalgudi, Raghunath V.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Sherman, Bradley K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
; FILE REFERENCE: PL-0017 US  
; CURRENT APPLICATION NUMBER: US/09/313.294A  
; CURRENT FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 7600  
; SOFTWARE: PERL Program  
; SEQ ID NO 1459  
; LENGTH: 289  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6476212 700550983H1  
US-09-313-294A-1459

Alignment Scores:  
Pred. No.: 1.05e-36 Length: 289  
Score: 368.00 Matches: 70  
Percent Similarity: 86.17% Conservative: 11  
Best Local Similarity: 74.47% Mismatches: 11  
Query Match: 69.43% Indels: 2  
DB: 4 Gaps: 1

US-10-629-953-4 (1-101) x US-09-313-294A-1459 (1-289)

Qy 1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGly-----GlyGluProAlaAlaLys 18  
Db 6 ATGTGGGGGAAGAAAGAGCGTTTCGGCGGTGGCGCAGCGGTGAGCCGCGCGCAAG 65

[illegible]

## RESULT 2

```

US-09-313-294A-1612
; Sequence 1612, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 1612
; LENGTH: 263
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700551220H1
; NAME/KEY: unsure
; LOCATION: 25, 181
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-1612

```





; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 18624  
; LENGTH: 479  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-621-976-18624

Alignment Scores:  
Pred. No.: 0.138 Length: 479  
Score: 84.00 Matches: 21  
Percent Similarity: 55.17% Conservative: 11  
Best Local Similarity: 36.21% Mismatches: 16  
Query Match: 15.85% Indels: 10  
DB: 4 Gaps: 2

US-10-629-953-4 (1-101) x US-09-621-976-18624 (1-479)

Qy 14 GluProAlaLysArgAlaAlaGlyAspGly----- 26  
Db 167 GAAACCTGTAAGAAACAAAGACAGGTGAGACTTCGAGAGCCCTGTCTCTTAA 226  
Qy 27 ---ProSerGluSerAlaAspAspAlaValAlaGlnileSerLysAsnArg 45  
Db 227 CAGAGCAGCAGCAGCAGATGATGATGTT-----CAGATTGGGAAATGAGGTAC 280  
Qy 46 ValAlaValArgThrTrpAsnGlyLysValValValAspAlaArgGluPheTyr 63  
Db 281 GTTAGTCTCGCGATTTTAAAGGCAAGTGCTAATTGATATTAGAGATATTGG 334

## RESULT 11

US-09-724-797-85  
; Sequence 85, Application US/09724797  
; Patent No. 6733998

## GENERAL INFORMATION:

; APPLICANT: Jon S. THORSON  
; TITLE OF INVENTION: MICROMONOPORA ECHINOSPORA GENES  
; TITLE OF INVENTION: ENCODING FOR BIOSYNTHESIS OF  
; TITLE OF INVENTION: CALICHEAMICIN AND SELF-RESISTANCE THERETO  
; FILE REFERENCE: 2653-40  
; CURRENT APPLICATION NUMBER: US/09/724,797  
; CURRENT FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 60/111,325  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 85  
; LENGTH: 1320  
; TYPE: DNA  
; ORGANISM:  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1320)  
US-09-724-797-85

Alignment Scores:  
Pred. No.: 1.88 Length: 1320  
Score: 79.50 Matches: 26  
Percent Similarity: 43.68% Conservative: 12  
Best Local Similarity: 29.89% Mismatches: 44  
Query Match: 15.00% Indels: 5  
DB: 4 Gaps: 2

US-10-629-953-4 (1-101) x US-09-724-797-85 (1-1320)

Qy 4 LysGlyAsnLysArgPheGlyGlyGluProAlaLysArgAlaAlaGly 23  
:::|||||

Db 957 CGCGCGCGCGCGGGTGGCGCGGAGCGCGCGGAGCAAGGTCCGGGG 1016  
Qy 24 AspAspGlyProSerGluSerAlaAspAspAlaValAlaGlnileSerLysAsn 43  
Db 1017 CGCGTGGCGCTGCGCCACGCTGCTGGCGCGCGCTGCTGGCAGCGCGGTGCC 1076  
Qy 44 ArgArgValAlaValArg-----ThrTrpAsnGlyLysValValAlaAspAlaArg 60  
Db 1077 CGCGGACCTGCGCAGCGTGCACCGCGCGCGCGCTGCTGGCGCGCTGCGCG 1136  
Qy 61 GluPheTyrGluLysAspGlyLysThrLeuProGlyLysGlyLysGlnLeuProMet 80  
Db 1137 GAGTTCTC-----GACGGCGAGCGCTCACCGCGCGGTGCGAGGTGCCCGACCGCG 1190  
Qy 81 AspGlnTrpLysLeuArg 87  
Db 1191 AACTACTGGGAGCGCGCGG 1211

## RESULT 12

US-09-252-991A-15140/c  
; Sequence 15140, Application US/09252991A  
; Patent No. 6551795

## GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 15140  
; LENGTH: 2058  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-15140

## Alignment Scores:

Pred. No.: 12.1 Length: 2058  
Score: 75.00 Matches: 26  
Percent Similarity: 46.67% Conservative: 9  
Best Local Similarity: 34.67% Mismatches: 20  
Query Match: 14.15% Indels: 20  
DB: 4 Gaps: 4

US-10-629-953-4 (1-101) x US-09-252-991A-15140 (1-2058)

Qy 2 TrpArgLysGlyAsnLysArgPheGlyGly-----GlyGlyGlu----- 14  
Db 252 TTCGCGCGCGCGCTGCGCTTCGCGGAGCGCGCTGCGGAGCGCGAGCGCGCGAG 193  
Qy 15 -----ProAlaLysArgArgAlaAlaGlyAspGlyProSerGluSerAlaAsp 32  
Db 192 GAAGCGGAGCG 133  
Qy 33 AspAspAlaValAlaGlnileSerLysAsnArgValAlaAlaValArgThrTrpAsn 52  
Db 132 GATCAGTCTGTTTCATGTGCGGGTT-----GTTCAATGTTGCT 94

## RESULT 13

US-09-252-991A-14623  
; Sequence 14623, Application US/09252991A  
; Patent No. 6551795

## GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 14623

; LENGTH: 2211

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-14623

Alignment Scores:

Pred. No.: 13.3 Length: 2211  
Score: 75.00 Matches: 26  
Percent Similarity: 46.67% Conservative: 9  
Best Local Similarity: 34.67% Mismatches: 20  
Query Match: 14.15% Indels: 20  
DB: 4 Gaps: 4

US-10-629-953-4 (1-101) x US-09-252-991A-14623 (1-2211)

QY 2 TrpArgLysGlyAsnLysArgPheGlyGly-----GlyGlyGlu----- 14  
Db 1906 TTCGCGGGGGCGCTTCGCGTTCGCGGGAGCGCGCTGGCGGGGAGCGCGCGAG 1965

QY 15 -----ProAlaLysArgArgAlaAlaGlyAspGlyProSerGluSerAlaAsp 32

Db 1966 GAAGGCGAGGCG 2025

QY 33 AspAspIleValValAlaGlnIleSerLysAsnArgValAlaValArgThrTrpAsn 52

Db 2026 GATCAGGTCTGTCATGTGCGGTT-----GTTCAATGTTGCT 2064

QY 53 GlyLysValValAlaAspIleArgGluPheTyrGluLysAspGly 67

Db 2065 GGCGCACTCGTCGCC-----ATCACCAGACAGATGCG 2097

RESULT 14

US-09-919-172-40/c

; Sequence 40, Application US/09919172

; Patent No. 6673545

; GENERAL INFORMATION:

; APPLICANT: Paris, Mary

; APPLICANT: Turner, Christopher M.

; TITLE OF INVENTION: PROSTATE CANCER MARKERS

; FILE REFERENCE: PA-0036 US

; CURRENT APPLICATION NUMBER: US/09/919,172

; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: 60/222,469

; PRIOR FILING DATE: 2000-07-28

; NUMBER OF SEQ ID NOS: 102

; SOFTWARE: PERL Program

; SEQ ID NO 40

; LENGTH: 2712

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. 6673545 1798379CB1

US-09-919-172-40

Alignment Scores:

Pred. No.: 17.4 Length: 2712  
Score: 75.00 Matches: 31  
Percent Similarity: 37.19% Conservative: 14  
Best Local Similarity: 25.62% Mismatches: 37  
Query Match: 14.15% Indels: 39  
DB: 4 Gaps: 5

US-10-629-953-4 (1-101) x US-09-919-172-40 (1-2712)

QY 1 MetTrpArgLysGlyAsn---LysArgPheGlyGlyGly----- 13

Db 2292 ATCTGGAATATGGCAATGCAAGAGGCGATGTTGGGGGGGGAATGGGCATGGCTAGAG 2233

QY 14 -----GluProAlaAlaLysArgArg-----Ala 21

Db 2232 CACAAACCTAGACAGCACAGATTCGCGCCCGCAGAACACAGAGGCGTCATCGTCTTA 2173

QY 22 AlaGlyAspGly----- 26

Db 2172 ACTGGTTGGAGAGATCTTTCAAAAACCTGTATGTGACATCTTTTCCCTCAAAGTG 2113

QY 27 ProSerGluSerAlaAspAspIleValValAlaGlnIleSerLysAsnArgArgVal 46

Db 2112 CCATCAGATATAATGACAAATAATTAACTTCTGCCATTAGAGGTACAGACAGAGCT 2053

QY 47 AlaVal-----ArgThrTrpAsnGlyLysVal-ValValAspIleArgGluPh 62

Db 2052 TCTACATCTCGAGTGGGAGAGTCTGGGTTGGGAGAGTGGTGGTGGGGCTCATCT 1993

QY 62 eTyrGluLysAspGlyLysThrIleuProGlyArgLysGlyIleGlnLeuProMetAspGl 82

Db 1992 TTATCTAAAGAGAAGGTCACTGCAACCAAGCTTCACTGAGCTACCTTTTACAGA 1933

QY 82 n 82

Db 1932 G 1932

RESULT 15

US-09-489-039A-1584/c

; Sequence 1584, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 1584

; LENGTH: 528

; TYPE: DNA

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-1584

Alignment Scores:

Pred. No.: 2.29 Length: 528  
Score: 74.50 Matches: 28  
Percent Similarity: 37.76% Conservative: 9  
Best Local Similarity: 28.57% Mismatches: 34  
Query Match: 14.06% Indels: 27  
DB: 4 Gaps: 5

US-10-629-953-4 (1-101) x US-09-489-039A-1584 (1-528)

QY 5 GlyAsnLysArgPheGlyGlyGly-----GluProAlaAlaLysArgArgAla 21

Db 399 GCCACCGGCGGT 340

QY 22 AlaGlyAspGlyProSerGluSer-----AlaAsp 32

Db 339 GCGGGTCACTCGGTCCAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGT 280

QY 33 AspAspIleValValAlaGlnIle-----SerLysAsnArgArgValAlaValArgThr 50

Db 279 CAGCAGGCTGTGCGCGCGGAAATCATCATGTCCCGCAGCGCGCGCGCGCGCGCG 220

QY 51 TrpAsnGlyLysValValAspIleArgGluPheTyrGluLysAspGlyLysThrLeu 70

```

Db      219  -----GCAGCCCGCGCGGCCAACAGCAGCGCGCGCGCGACT 181
Qy      71  -----ProGlyArgGlyGlyIleGlnLeuProMetAspGln 82
Db      180  GGCGCGCGCGCAACCGCAGCGCGGAAATGGCCAGCATCAGCCAGATCAGCAG 127

```

Search completed: January 12, 2005, 03:29:36  
Job time : 101 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2005, 02:45:26 ; Search time 462 Seconds  
(without alignments)  
1254.369 Million cell updates/sec

Title: US-10-629-953-4

Perfect score: 530

Sequence: 1 MRKGNRRGGGEPAAKRR.....QWKILRDNIKAIDEAIKENA 101

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 4293498 seqs, 2868903791 residues

Total number of hits satisfying chosen parameters: 8586996

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool.h/US10629953/runat\_10012005\_172538\_9528/app\_query.fasta\_1.263  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -WAP=US10629953 @CGN 1 1 480 @runat\_10012005\_172538\_9528  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:  
18: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:  
19: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:  
20: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:  
21: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

\*

Result No.	Score	Query Match	Length	DB ID	Description
1	530	100.0	1754	17	US-10-437-963-83372
2	398	75.1	660	16	US-10-425-114-4500
3	395	74.5	757	18	US-10-425-115-74342
4	304.5	57.5	310	9	US-09-294-093B-67
5	269	50.8	663	16	US-10-424-599-38767
6	261.5	49.3	917	16	US-10-424-599-38768
c 7	183	34.5	496	10	US-09-770-961-520
8	182.5	34.4	794	16	US-10-425-114-20666
c 9	182.5	34.4	911	16	US-10-424-599-4356
10	182	34.3	1030	17	US-10-437-963-59601
c 11	179	33.8	1479	16	US-10-424-599-9157
c 12	177	33.4	497	16	US-10-424-599-127795
13	172	32.5	2776	13	US-10-087-192-1784
14	166.5	31.4	933	16	US-10-425-114-24336
15	166.5	31.4	935	16	US-10-425-114-33158
16	166.5	31.4	1225	18	US-10-425-115-154421
17	166	31.3	658	16	US-10-425-114-21115
18	166	31.3	658	18	US-10-425-115-20439
19	163	30.8	483	10	US-09-918-995-25317
20	160.5	30.3	384	15	US-10-177-293-349
21	160.5	30.3	384	15	US-10-257-021-21
22	160.5	30.3	384	18	US-10-370-715B-281
23	160.5	30.3	460	10	US-09-918-995-35177
24	160.5	30.3	466	13	US-10-087-192-1787
25	160.5	30.3	469	16	US-10-242-535A-39519
26	160.5	30.3	469	16	US-10-085-783A-39519
27	160.5	30.3	500	16	US-10-242-535A-51406
28	160.5	30.3	500	16	US-10-085-783A-51406
29	160.5	30.3	592	9	US-09-895-828-179
30	160.5	30.3	592	15	US-10-114-666-179
31	160.5	30.3	601	9	US-09-895-828-342
32	160.5	30.3	601	15	US-10-114-666-342
33	160.5	30.3	674	9	US-09-738-973-477
34	160.5	30.3	674	9	US-09-854-133-477
35	160.5	30.3	685	15	US-10-144-649A-477
36	160.5	30.3	685	15	US-10-242-535A-45670
37	160.5	30.3	685	16	US-10-085-783A-45670
38	160.5	30.3	689	9	US-09-920-345-29
39	160.5	30.3	1387	14	US-10-198-846-13428
40	160.5	30.3	1391	18	US-10-357-930-25730
41	160.5	30.3	3527	9	US-09-925-297-320
42	159.5	30.1	400	16	US-10-242-535A-51580
43	159.5	30.1	400	16	US-10-085-783A-51580
44	157.5	29.7	432	10	US-09-918-995-6997
45	150.5	28.4	420	16	US-10-242-535A-52202

ALIGNMENTS

RESULT 1  
US-10-437-963-83372  
; Sequence 83372, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 83372  
; LENGTH: 1754

; TYPE: DNA  
 ; ORGANISM: Oryza sativa  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_8270C.1  
 US-10-437-963-83372

Alignment Scores:  
 Pred. No.: 3.36e-57 Length: 1754  
 Score: 530.00 Matches: 101  
 Percent Similarity: 100.00% Conservativity: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 17 Gaps: 0

US-10-629-953-4 (1-101) x US-10-437-963-83372 (1-1754)

QY 1 MetTrrpArgLysGlyAsnLysArgPheGlyGlyGlyGluProAlaLysArgArg 20  
 Db 1123 ATGTGGCGGAAGGGAACAAGCGGTTCCGGCGCGCGCGGAGCGCGCGCGCGT 1182  
 QY 21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspLysValValAlaGlnIle 40  
 Db 1183 GCCGCGCGGACGACGCGCCCTCCGAGAGCGCGACGACGATATCGTCGCGCCGAGATA 1242  
 QY 41 SerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValAspLysArg 60  
 Db 1243 TCGAAGACAGAGGGTGGCGGTTCGGACCTGGACGCGCGCGCGCGCGCGT 1302  
 QY 61 GluPheTrrpLysAspGlyLysThrLeuProGlyArgLysGlyGlnLeuProMet 80  
 Db 1303 GAGTCTTACGAGAAGGACGCGACCGCTCCCGCGCGCGCAAGGATATACAGCTCCGCGATG 1362  
 QY 81 AspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100  
 Db 1363 GATCAGTGAAGATACCTGAGGACAATATCAAGCTATAGTATAGGCGCATCAAGGAGAAAT 1422  
 QY 101 Ala 101  
 Db 1423 GCG 1425

RESULT 2

US-10-425-114-4500.  
 ; Sequence 4500, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lin, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E.  
 ; APPLICANT: Tabaska, Jack E.  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 4500  
 ; LENGTH: 660  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: 700380835\_FLI

Alignment Scores:  
 Pred. No.: 5.65e-41 Length: 660  
 Score: 398.00 Matches: 76  
 Percent Similarity: 87.25% Conservativity: 13  
 Best Local Similarity: 74.51% Mismatches: 11  
 Query Match: 75.09% Indels: 2  
 DB: 16 Gaps: 1

US-10-629-953-4 (1-101) x US-10-425-114-4500 (1-660)

QY 1 MetTrrpArgLysGlyAsnLysArgPheGlyGlyGlyGluProAlaLys 18  
 Db 51 ATGTGGCGGAAGGGAACAAGCGGTTCCGGCGGTGGCGGAGCGGTGAGCCCGCGGCAAG 110  
 QY 19 ArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspLysValValAla 38  
 Db 111 CGCCAGCGCGCGAGGAGCGATGCGCCCTCCGAATCGCTGAAGACGCTACCGTCGTAGCC 170  
 QY 39 GlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValAsp 58  
 Db 171 GAGATATCGAAGAACAAAGAGGTGTCCGTTAGGAGCTGGAAGCGCGAGGCTTTCGTCGAC 230  
 QY 59 IleArgGluPheTrrpLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeu 78  
 Db 231 TTACGCGAGTTCTACTTCAAGGACGCAAGACTCTCCACCGCAAGGATATATCACTC 290  
 QY 79 ProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLys 98  
 Db 291 CAATTAGATCAGTGGGAAGATATTGAAGGACAACATCAAGCCATAAATGAGGCAATCGAG 350  
 QY 99 GluAsn 100  
 Db 351 GAAAT 356

RESULT 3

US-10-425-115-74342  
 ; Sequence 74342, Application US/10425115  
 ; Publication No. US20040214272A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(53222)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,115  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 369326  
 ; SEQ ID NO 74342  
 ; LENGTH: 757  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: MRT4577\_167799C.1

Alignment Scores:

Pred. No.: 1.59e-40 Length: 757  
 Score: 395.00 Matches: 75  
 Percent Similarity: 87.25% Conservativity: 14  
 Best Local Similarity: 73.53% Mismatches: 11  
 Query Match: 74.53% Indels: 2  
 DB: 18 Gaps: 1

US-10-629-953-4 (1-101) x US-10-425-115-74342 (1-757)

QY 1 MetTrrpArgLysGlyAsnLysArgPheGlyGlyGlyGluProAlaLys 18  
 Db 72 ATGTGGCGGAGGGAAGAAGCGTTTCGGCGGTGGCGGAGCGGTGAGCCCGCGGCAAG 131  
 QY 19 ArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspLysValValAla 38  
 Db 132 CGCCAGCGCGGAGGAGCGATGCGCCCTCCGAATCGCTGAAGACGCTACCGTCGTAGCC 191  
 QY 39 GlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValAsp 58  
 Db 192 GAGATATCGAAGAACAAAGAGGTGTCCGTTAGGAGCTGGAAGCGCGAGGCTTTCGTCGAC 251  
 QY 59 IleArgGluPheTrrpLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeu 78  
 Db 591 GAGATATCGAAGAACAAAGAGGTGTCCGTTAGGAGCTGGAAGCGCGAGGCTTTCGTCGAC 230

Db 252 TTACGGAGTCTTACTTCAAGGACGGCAAGACTCTCCCAACCGCAAGGATATATCACTC 311  
Qy 79 ProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLys 98  
Db 312 CAATTAGATCAGTGGAGATATTGAAGACAACATCAAGCCATAAATGAGGCCATCGAG 371  
Qy 99 GluAsn 100  
Db 372 GAAAT 377

RESULT 4  
US-09-294-093B-67  
; Sequence 67, Application US/09294093B  
; Patent No. US20010051335A1  
; GENERAL INFORMATION:  
; APPLICANT: Lalugudi, Raghunath, V.  
; APPLICANT: Ito, Laura, Y.  
; APPLICANT: Sherman, Bradley, K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL  
; FILE REFERENCE: PL-0009 US  
; CURRENT APPLICATION NUMBER: US/09/294,093B  
; PRIOR FILING DATE: 1999-04-16  
; PRIOR APPLICATION NUMBER: 60/082,567  
; PRIOR FILING DATE: April 21, 1998  
; NUMBER OF SEQ ID NOS: 6207  
; SOFTWARE: PERL Program  
; SEQ ID NO 67  
; LENGTH: 310  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700281907H2  
; NAME/KEY: unsure  
; LOCATION: 2, 30, 295  
; OTHER INFORMATION: a, t, c, g, or other  
US-09-294-093B-67

Alignment Scores:  
Pred. No.: 1.64e-29 Length: 310  
Score: 304.50 Matches: 65  
Percent Similarity: 81.44% Conservative: 14  
Best Local Similarity: 67.01% Mismatches: 13  
Query Match: 57.45% Indels: 5  
DB: 9 Gaps: 2

US-10-629-953-4 (1-101) x US-09-294-093B-67 (1-310)  
Qy 1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGly-----GlyGluProAlaAlaLys 18  
Db 26 ATGTNGGGGGGGAAGAAAGCGTTTCGGCGGTGGCGGCGAGCGCGCGCGGCGCAAG 85  
Qy 19 ArgArgAlaAla-GlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAl 38  
Db 86 CGCAGCGCGCGGAGGAGACGATGCGCCCTCCGAATCCGCTGAAGACGGTACCGCTCGTAGC 145  
Qy 38 aGlnIleSerLysAsnArgValAlaValArgThrTrpAsnGlyLysValValValas 58  
Db 146 CGAGATATCGAAGACAAGAGGTATCCGTAGAGCTGGAAGCGAGGCTCTTCGTGCA 205  
Qy 58 pIleArgGluPheTyrcGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGln-L 78  
Db 206 CTTACGCGAGTTCTACTTCAAGGACGCGCAAGACTCTCCCGCCGCAAA---GTATATCA 262  
Qy 78 euProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIle 93  
Db 263 TTCCAATTAGATCATGGAAGATATTGAAGGACNACATCAAGGCCATA 309

RESULT 5  
US-10-424-599-38767  
; Sequence 38767, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 38767  
; LENGTH: 917  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_135007C.1

; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 38767  
; LENGTH: 663  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_135006C.1  
US-10-424-599-38767

Alignment Scores:  
Pred. No.: 1.24e-24 Length: 663  
Score: 269.00 Matches: 56  
Percent Similarity: 66.99% Conservative: 13  
Best Local Similarity: 54.37% Mismatches: 14  
Query Match: 50.75% Indels: 20  
DB: 16 Gaps: 3

US-10-629-953-4 (1-101) x US-10-424-599-38767 (1-663)  
Qy 17 AlaLysArgArgAlaAlaGlyAspAspGlyProSerGluSer----- 30  
Db 178 GCGAAGAGAGA-----GATGACGACGGTCTTCCGACGCCGACTCTGAAGGCCAGCG 231  
Qy 31 -----AlaAspAspAsp-----IleValValAla 38  
Db 232 CGCCCAAGAGTCCCTCAAGAAGGATTCCGATGACGATCCCGACTATTACGTTTGC 291  
Qy 39 GlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValasp 58  
Db 292 GAGATTTTCAAGAAACAGAGGGGTTCGCGTGAAGAACTGGAAGAGCAGCATTTATGTTGAC 351  
Qy 59 IleArgGluPheTyrcGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeu 78  
Db 352 ATTCCGCGAGTTTACGTCAAAGATGCGAAGCAATTCCTGCGCAAAAGGTATCTCTTTG 411  
Qy 79 ProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLys 98  
Db 412 ACCATGATCATGTGGAATGCTTCTGTAATCATGTTGAAGAAATTCACAGGCAATTAAT 471  
Qy 99 GluAsnAla 101  
Db 472 GAGAAATCT 480

RESULT 6  
US-10-424-599-38768  
; Sequence 38768, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 38768  
; LENGTH: 917  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_135007C.1

US-10-424-599-38768

Alignment Scores:  
 Pred. No.: 1,61e-23 Length: 917  
 Score: 261.50 Matches: 48  
 Percent Similarity: 68.13% Conservative: 14  
 Best Local Similarity: 52.75% Mismatches: 24  
 Query Match: 49.34% Indels: 5  
 DB: 16 Gaps: 1

US-10-629-953-4 (1-101) x US-10-424-599-38768 (1-917)

QY 11 GlyGlyGlyGluProAlaAlaLysArgAlaAlaGlyAspAspGlyProSerGluSer 30  
 DB 305 GGCCACGCGCGCCATGAAGTCCCTCAAGAAGGATTCGACGACGATCCC----- 355  
 QY 31 AlaAspAspAspLeValValAlaGlnIleSerLysAsnArgValAlaValArgThr 50  
 DB 356 -----GACTCTGTTACCATTTTCGAAATTTTCGAAGAACAGGAGGTTGCCGCTGAGGAAC 409  
 QY 51 TrpAsnGlyLysValValAlaAspIleArgGluPheTyrGluLysAspGlyLysThrLeu 70  
 DB 410 TGAAGGCGAGCATTAATGTTGATTCGCGAGTTTACGTCAAGATGCGCAAGCAATTG 469  
 QY 71 ProGlyArgLysGlyIleGlnLeuProMetAspGlnTrpLysIleLeuArgAspAsnIle 90  
 DB 470 CCTGGCAGGAAGGTATCTCGTTGACCATGTCAGTGGAAATGTGCTTCGTAATCATGTT 529  
 QY 91 LysAlaIleAspGluAlaIleLysGluAsnAla 101  
 DB 530 GGAAGAATTTGACAAAGGCGAGTGAATGATTC 562

RESULT 7

US-09-770-961-520/c  
 ; Sequence 520, Application US/09770961  
 ; Publication No. US20030115639A1

; GENERAL INFORMATION:  
 ; APPLICANT: Gorlach, Jorn

; APPLICANT: An, Yong-Qiang  
 ; APPLICANT: Hamilton, Carol M.

; APPLICANT: Price, Jennifer L.  
 ; APPLICANT: Raines, Tracy M.

; APPLICANT: Yu, Yang  
 ; APPLICANT: Rameaka, Joshua G.

; APPLICANT: Page, Amy  
 ; APPLICANT: Matthew, Abraham V.

; APPLICANT: Ledford, Brooke L.  
 ; APPLICANT: Woessner, Jeffrey P.

; APPLICANT: Haas, William David  
 ; APPLICANT: Garcia, Carlos A.

; APPLICANT: Krickler, Maja  
 ; APPLICANT: Slader, Ted

; APPLICANT: Davis, Keith R.  
 ; APPLICANT: Allen, Keith

; APPLICANT: Hoffman, Neil  
 ; APPLICANT: Hurlan, Patrick

; TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
 ; FILE REFERENCE: 2026 (PARA-015PRV)

; CURRENT APPLICATION NUMBER: US/09/770,961  
 ; CURRENT FILING DATE: 2001-01-26

; PRIOR APPLICATION NUMBER: 60/178,466  
 ; PRIOR FILING DATE: 2000-01-27

; NUMBER OF SEQ ID NOS: 999  
 ; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 520  
 ; LENGTH: 496

; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana

US-09-770-961-520

Alignment Scores:  
 Pred. No.: 6.89e-14 Length: 496

Score: 183.00 Matches: 37  
 Percent Similarity: 56.25% Conservative: 17  
 Best Local Similarity: 38.54% Mismatches: 22  
 Query Match: 34.53% Indels: 20  
 DB: 10 Gaps: 2

US-10-629-953-4 (1-101) x US-09-770-961-520 (1-496)

QY 4 LysGlyAsnLysArgPheGlyGlyGlyGluProAlaAlaLysArgAlaAlaGly 23  
 DB 459 AAAGGAACAAGAGTTT----- 442  
 QY 24 AspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLysAsn 43  
 DB 441 GATGAT-----GACGGCATCTTATCATTTGCGATTATCGGATAAG 400  
 QY 44 ArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArgGluPheTyr 63  
 DB 399 AGAAGAGTGCAGGATTCAGGAATTTAAAGGGAAGAGTTTGGTTTCTATCAGAGATATTAC 340  
 QY 64 GluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMetAspGlnTrp 83  
 DB 339 AGAAGAGTGGCAAGAACTTCTACTTCTAAAGGAATAGCTTAACAGATGAACATGG 280  
 QY 84 LysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99  
 DB 279 TCAACCTTCAAGAAACATGCCAGCCATCGAAATGCTGTCAAGAA 232

RESULT 8

US-10-425-114-20666

; Sequence 20666, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E.

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 20666

; LENGTH: 794

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB3242-699-H3\_FLI

US-10-425-114-20666

Alignment Scores:

Pred. No.: 1.38e-13 Length: 794

Score: 182.50 Matches: 36

Percent Similarity: 63.64% Conservative: 20

Best Local Similarity: 40.91% Mismatches: 23

Query Match: 34.43% Indels: 9

DB: 16 Gaps: 2

US-10-629-953-4 (1-101) x US-10-425-114-20666 (1-794)

QY 12 GlyGlyGluProAlaAlaLysArgAlaAlaGlyAspGlyProSerGluSerAla 31

DB 340 GGAGGA-----GCTTCCAAAGGCAAGGATGATGATGAAGGC----- 378

QY 32 AspAspAspIleValValAlaGlnIleSerLysAsnArgValAlaValArgThrTrp 51

DB 379 -----GATCTCATCATCTGCGGCTTTCAGATAAGAGGCGTTCAGGATTCAGGATTC 432

QY 52 AsnGlyLysValValAspIleArgGluPheTyrGluLysAspGlyLysThrLeuPro 71

```
Db 433 AGAGGGAAACATGGTCTCCATTCGGGAGTATTATATAAAGAGTGCAGAGAACTTCCT 492
Qy 72 GlyArgLysGlyLeuGlnLeuProMetAspGlnTrpLysLeuArgAspAsnIleLys 91
Db 493 ACTTCCAAAGGAATAAGTTTGACAGAAGCAGTGCGTTCAGCCTTTTAAGAAAATGTGCCT 552
Qy 92 AlaIleAspGluAlaIleLysGlu 99
Db 553 GCCATAGAAAAGCCATTAAAGAA 576

RESULT 9
US-10-424-599-4356
; Sequence 4356, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 4356
; LENGTH: 911
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_10393C.1
US-10-424-599-4356
Alignment Scores:
Pred. No.: 1,61e-13 Length: 911
Score: 182.50 Matches: 36
Percent Similarity: 63.64% Conservative: 20
Best Local Similarity: 40.91% Mismatches: 23
Query Match: 34.43% Indels: 9
DB: 16 Gaps: 2

US-10-629-953-4 (1-101) x US-10-424-599-4356 (1-911)
Qy 12 GlyGlyGluProAlaAlaLysArgAlaGlyAspAspGlyProSerGluSerAla 31
Db 410 GGAGGA-----GTTCCAAAGGCGAAGCAGTACGATGATGAAGC----- 448
Qy 32 AspAspIleValValAlaGlnIleSerLysAsnArgArgValAlaValArgThrTrp 51
Db 449 -----GATCTCATCTCTCGAGCTTTTCAGATAAGAGAAGGTTGACGATTCAGGATTTC 502
Qy 52 AsnGlyLysValValValAspIleArgGluPheTyrGluLysAspGlyLysThrLeuPro 71
Db 503 AGAGGGAAACATATGTTGTCATTCGGGAGTATTATATAAAGAGTGCAGAGAACTTCCT 562
Qy 72 GlyArgLysGlyLeuGlnLeuProMetAspGlnTrpLysLeuArgAspAsnIleLys 91
Db 563 ACTTCCAAAGGAATAAGTTTGACAGAAGCAGTGCGTTCAGCCTTTTAAGAAAATGTGCCT 622
Qy 92 AlaIleAspGluAlaIleLysGlu 99
Db 623 GCCATAGAAAAGCCATTAAAGAA 646

RESULT 10
US-10-437-963-59601/C
; Sequence 59601, Application US/10437963
; Publication No. US2004012334JAI
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
```

```
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 59601
; LENGTH: 1030
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_61208C.1
US-10-437-963-59601
Alignment Scores:
Pred. No.: 2,15e-13 Length: 1030
Score: 182.00 Matches: 41
Percent Similarity: 62.00% Conservative: 21
Best Local Similarity: 41.00% Mismatches: 27
Query Match: 34.34% Indels: 11
DB: 17 Gaps: 3

US-10-629-953-4 (1-101) x US-10-437-963-59601 (1-1030)
Qy 3 ArgLysGlyAsnLysArgPheGlyGlyGlyGluPro-----AlaAlaLysArgAr 20
Db 770 CGAGGAGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 711
Qy 20 gAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspIleValValAlaGlnIle 40
Db 710 GGAGTAGCATGACGAGGGA-----GACCTCATCTCTCGCGCT 672
Qy 40 eSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleAr 60
Db 671 TTCGGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 612
Qy 60 gGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyLysGlnLeuProMe 80
Db 611 AGAGTACTACTTCAGGACGCGCAGGAGCTGCCCGCC-----AAGGGATAGTTTGACAGT 555
Qy 80 tAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99
Db 554 TGAGCAGTGGGAGCATTCGTGATTCTGTGTGCCAGCAATAGAGGATGCCATTAAAAAG 497

RESULT 11
US-10-424-599-9157
; Sequence 9157, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 9157
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_108279C.1
US-10-424-599-9157
Alignment Scores:
Pred. No.: 7,85e-13 Length: 1479
Score: 179.00 Matches: 38
```

Percent Similarity:	63.53%	Conservative:	16
Best Local Similarity:	44.71%	Mismatches:	28
Query Match:	33.77%	Indels:	3
DB:	16	Gaps:	1

US-10-629-953-4 (1-101) x US-10-424-599-9157 (1-1479)

Qy	18	LysArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAsp-AspAsp-----II	35
Db	661	CGAAGAAGAGTCCCGAGAGACAGAGAGTCTGCCCAAACTGAAGAGGATGATCCGCAACG	720
Qy	35	evalValAlaGlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysVa	55
Db	721	CGTATTTCACCTGTCCACAGAGGNAAGTGGCGGTGAAAGATTTCAAAGGGCAAC	780
Qy	55	lValValAspIleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGI	75
Db	781	CCTGGTCTCAATTAGGGAGTCTATATGAAAGATGAAAAACCACTTCCTGTTTCGAAGG	840
Qy	75	YleGlnLeuProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGI	95
Db	841	GATAGTATTATCTTCGAACAATGTCTGCACCTTCAGAGAGAGTGTCTCTGCCATAGAGGA	900
Qy	95	uAlaIleLysGlu 99	
Db	901	AGCTATCAAAAAG 913	

**RESULT 12**

US-10-424-599-127795/c  
; Sequence 127795, Application US/10424599  
; Publication No. US20040031072A1

```

; GENERAL INFORMATION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
;
; APPLICANT: Cao Yongwei
; APPLICANT: Zhou Yihua
; APPLICANT: Kovalic David K
; APPLICANT: La Rosa Thomas J

```

OTHER INFORMATION: Clone ID: PAT\_MRT3847\_86402C.1  
US-10-424-599-127795

Alignment Scores:	
Pred. No.:	3.97e-13
Score:	177.00
Percent Similarity:	66.22%
Best Local Similarity:	41.89%
Query Match:	33.40%
DB:	16
US-10-629-953-4	(1-101)
x US-10-424-599-127795	(1-497)

```
Db      312 TTTAGAAAATGTGCTCCCATAGAAAGCATTAAAGAA 271
          ::: |||::: |||||:::|||||:::|||||:::
RESULT 13
US-10-087-192-1784
; Sequence 1784, Application US/10087192
; Publication No. US20020182586A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: fastseq for windows Version 4.0
; SEQ ID NO 1784
; LENGTH: 2776
;

```

; ORGANISM: Mus  
US-10-087-192-1784

Alignment Scores:		
Pred. No.:	1.25e-11	2776
Score:	172.00	36
Percent Similarity:	64.37%	Conservative: 20
Best Local Similarity:	41.38%	Mismatches: 29
Query Match:	32.45%	Indels: 2
DB:	13	Gaps: 2

US-10-629-953-4 (1-101) X US-10-087-192-1784 (1-2776)

Qy	14	GluProAlaAlaLysArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAsp	33
Db	163	AAAGCTGTGTAGACACTTTCTAGACAGCTGCATCTCTCAAGCAGACAGCAGCAGCAGAGAT	222
Qy	34	AspIleValValAlaGlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGly	53
Db	223	GACACATGTTCC--CAGATTGGAAAGATGAGATATGTTCAGTGTTCGGGACTTCAAAGGA	279
Qy	54	LysValValValAspIleArgGlnPheTyr---GluLysAspGlyLysThrLeuProGly	72
Db	280	AAAATTCCTAATTGATATTAGAGAAATATTGGATGAGTTCAGAAAGGTGAAATCAAAACCCAGGA	339
Qy	73	ArgLysGlyIleGlnLeuProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAla	92
Db	340	AGAAAGGTATTCTTTTAAACATGACACATGGAGCCAGCTCAAGACAGATCTCTCGAT	399
Qy	93	IleAspGluAlaIleLysGlu	99
Db	400	ATAGATGACCCAGTAAGAAAG	420

**RESULT 14**

US-10-425-114-24336  
; Sequence 24336, Application US/10425114  
; Publication No. US20040034888A1

```

; GENERAL INFORMATION:
; APPLICANT: Liu, Jingsong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaeka, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21 (5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28

```

DQ		365	AGAGTACGATGACCAAGGGTGACCCTAATCCCTGTGTCGCCCTCCTCGAGGCACAAAGAGAGAGGGTGAC	424
Qy	47	aValArgThrTrpAsnGlyLysValValValAspIleAargGluPheTyrGluLuYaAspGl	67	

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2005, 01:19:56 ; Search time 2548 Seconds  
(without alignments)

1444.432 Million cell updates/sec

Title: US-10-629-953-4

Perfect score: 530

Sequence: 1 MWRKGNKRFGGGEPAAKRR.....QWKILRDNKAIDEAIKENA 101

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgm2/USPTO.spool.h/US10629953/runat\_10012005\_172536\_9425/app.query.fasta\_1.263  
-DB=EST -QWMT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=45  
-OUTALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USBR=US10629953 @CGN 1 1 3437 @runat\_10012005\_172536\_9425 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	530	100.0	697	6	CB684473
2	530	100.0	784	6	CB684474
3	517	97.5	693	1	AU082888
4	484	91.3	530	7	D15350
5	425	80.2	523	6	CA116904
6	425	80.2	631	6	CA151445
7	425	80.2	635	6	CA230818
8	425	80.2	637	6	CA259070
9	425	80.2	649	6	CA189554

10	416	78.5	440	7	CF760736
11	415.5	78.4	564	6	CD863080
12	415	78.3	855	6	CA129813
13	412.5	77.8	532	5	BQ838824
14	409.5	77.3	579	5	BU998906
15	409.5	77.3	618	6	CB860066
16	409.5	77.3	618	6	CB860077
17	409.5	77.3	619	5	BU998918
18	409.5	77.3	1072	7	CK212042
19	400.5	75.6	539	1	AU092741
20	398	75.1	568	6	CB250145
21	398	75.1	596	6	CD568912
22	396	74.7	470	6	CD913769
23	395	74.5	603	1	AI691445
24	392	74.0	606	6	CA189705
25	391	73.8	403	2	BE424113
26	391	73.8	618	1	AI734589
27	390	73.6	566	2	BE186543
28	389.5	73.5	460	6	CA710341
29	381	71.9	623	7	CF635823
30	375	70.8	516	6	CF019517
31	373.5	70.5	586	6	CD568913
32	363	68.5	588	7	CF627159
33	363	68.5	588	7	CF627159
34	363	68.5	652	3	AY112536
35	362.5	68.4	443	6	C28075
36	347.5	65.6	396	6	C27852
37	338.5	63.9	555	4	BI795760
38	329	62.1	288	1	AU068341
39	329	62.1	542	5	BU498860
40	299	56.4	485	1	AI943769
41	272.5	51.4	586	6	CB079577
42	272	51.3	551	6	CB922664
43	272	51.3	570	5	BQ796180
44	272	51.3	601	6	CB339739
45	272	51.3	620	6	CB921948

#### ALIGNMENTS

RESULT 1  
CB684473

LOCUS

DEFINITION

clone OSJNEF14A12 5', mRNA sequence.

CB684473

CB684473.1

GI:29688198

EST.

ORYZA SATIVA (JAPONICA CULTIVAR-GROUP)

ORYZA SATIVA (JAPONICA CULTIVAR-GROUP)

EUKARYOTA; VIRIDIPLANTAE; STREPTOPHYTA;

SPERMATOPHYTES; MAGNOLIOPHYTES; LILIOPSIDA; POALES; POACEAE;

HRHARTOIDEAE; ORYZACEAE; ORYZA.

1 (bases 1 to 697)

Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,

Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.

Large-scale identification of ESTs involved in the interaction

between rice and Magnaporthe grisea

Unpublished (2003)

Contact: Rod Wing

Arizona Genomics Institute

University of Arizona

Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ

85721-0088, USA

Tel: 520 626 3967

Fax: 520 621 9288

Email: http://genome.arizona.edu

PCR Primers

FORWARD: gta aac cgg cca gtc

BACKWARD: gga aac agc tat gac cat g

Plate: 14 Row: A Column: 12

Seq primer: gta aac cga cgg cca gtc.

Location/Qualifiers

FEATURES

```

source
1. .697
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEF14A12"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEF"
/notes="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2: XhoI; Uninfected Control"

ORIGIN
Alignment Scores:
Pred. No.: 3.1e-52 Length: 697
Score: 530.00 Matches: 101
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-629-953-4 (1-101) x CB684473 (1-697)
Qy 1 MetTTPArgLySGLyAenLySArgPheGlyGlyGluProAlaLalysArgArg 20
Db 55 ATGTGGCGGAGGGAACAGCGTTTCGGCGGCGCGGAGCGCGGCGCGCGCGT 114
Qy 21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspLeValValAlaGlnIle 40
Db 115 GCGCGCGGAGGAGCGCGCGCTCCGAGAGCGCGCGACGACGATATCGTCTCGCCAGATA 174
Qy 41 SerLySAsnArgArgValAlaValArgThrTTPAsnGlyLySValValAlaAspIleArg 60
Db 175 TCGAAGAACAGGAGGTTGGCGGTTCGAGACCTTGAACGCGCAAGGTCTGCTCGACATCCG 234
Qy 61 GluPheTyRGLyLySAspGlyLySThrLeuProGlyArgLySGLyIleGlnLeuPromet 80
Db 235 GAGTTCTACGAGAGGAGCGAGACCGCTCCCGCGCGCGCAAGGTATACAGCTCCCGATG 294
Qy 81 AspGlnTTPlySLeuArgAspAsnIleLySAlaIleAspGluAlaIleLySGLuAsn 100
Db 295 GATCAGTGAAGATATCTGAGGACATATCAAGCTATAGATGAGGCGCATCAAGGAGAAT 354
Qy 101 Ala 101
Db 355 GCG 357

RESULT 2
CB684474/c
LOCUS OSJNEF14A12.1 784 bp mRNA linear EST 09-APR-2003
DEFINITION OSJNEF14A12.1 Oryza sativa (japonica cultivar-group) cDNA
clone OSJNEF14A12.1, mRNA sequence.
ACCESSION CB684474
VERSION CB684474.1 GI:29688199
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 784)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967

Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gta g
BACKWARD: gga aac agc tat gac cat g
Plate: 14 row: A column: 12
Seq primer: gga aac agc tat gac cat g.
FEATURES
Location/Qualifiers
source
1. .784
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEF14A12"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEF"
/notes="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2: XhoI; Uninfected Control"

ORIGIN
Alignment Scores:
Pred. No.: 3.57e-52 Length: 784
Score: 530.00 Matches: 101
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-629-953-4 (1-101) x CB684474 (1-784)
Qy 1 MetTTPArgLySGLyAenLySArgPheGlyGlyGluProAlaLalysArgArg 20
Db 640 ATGTGGCGGAGGGAACAGCGTTTCGGCGGCGCGGAGCGCGGCGCGCGT 581
Qy 21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspLeValValAlaGlnIle 40
Db 580 GCGCGCGGAGGAGCGCGCGCTCCGAGAGCGCGCGACGATATCGTCTCGCCAGATA 521
Qy 41 SerLySAsnArgArgValAlaValArgThrTTPAsnGlyLySValValAlaAspIleArg 60
Db 520 TCGAAGAACAGGAGGTTGGCGGTTCGAGACCTTGAACGCGCAAGGTCTGCTCGACATCCG 461
Qy 61 GluPheTyRGLyLySAspGlyLySThrLeuProGlyArgLySGLyIleGlnLeuPromet 80
Db 460 GAGTTTTCAGAGAGGAGCGCAAGACCTCCCGCGCGCGCAAGGTATACAGCTCCCGATG 401
Qy 81 AspGlnTTPlySLeuArgAspAsnIleLySAlaIleAspGluAlaIleLySGLuAsn 100
Db 400 GATCAGTGAAGATATCTGAGGACATATCAAGCTATAGATGAGGCGCATCAAGGAGAAT 341
Qy 101 Ala 101
Db 340 GCG 338

RESULT 3
AU082888
LOCUS AU082888 Rice callus Oryza sativa (japonica cultivar-group) cDNA
DEFINITION clone C53889, mRNA sequence.
ACCESSION AU082888
VERSION AU082888.1 GI:7212585
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 693)
AUTHORS Sasaki,T. and Yamamoto,K.
TITLE Rice cDNA from callus (2000)
JOURNAL Unpublished (2000)

```

COMMENT Contact: Takuji Sasaki  
National Institute of Agrobiological Resources  
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
305-8602, Japan  
Tel: 81-298-38-7441  
Fax: 81-298-38-7468  
Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
PROJECT = RGP

FEATURES Location/Qualifiers  
source  
1. .693  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="CS3889"  
/clone\_lib="Rice callus"  
/note="Vector: pBluescript II SK+; Site 1: SalI; Site 2:  
NotI; cDNA prepared from rice callus mRNAs by using  
oligo(dT) as a primer and ligating to the SalI-NotI site  
of pBluescript II SK+ phagemid."

ORIGIN  
Alignment Scores:  
Pred. No.: 1.05e-50 Length: 693  
Score: 517.00 Matches: 100  
Percent Similarity: 99.01% Conservatives: 0  
Best Local Similarity: 99.01% Mismatches: 1  
Query Match: 97.55% Indels: 0  
DB: 1 Gaps: 0

US-10-629-953-4 (1-101) x AU082888 (1-693)

Qy 1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGlyGluProAlaAlaLysArgArg 20  
|||  
Db 62 ATGTGNCGGAAGCGAACAAGCGGTTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGT 121  
|||  
Qy 21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAlleValAlaGlnIle 40  
|||  
Db 122 GCCCGCGGGGACGACGGCGCTCCGAGAGCGCCGACGATATCGTCGCGCCAGATA 181  
|||  
Qy 41 SerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValAlaAspIleArg 60  
|||  
Db 182 TCGAAGAACAGAGGGTGGCGGTGGGACCTGGAAACGGCAGAGTGTGTCGACATCGC 241  
|||  
Qy 61 GluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMet 80  
|||  
Db 242 GAGTCTACGAGAGGACGCGCAAGACCTCCCGCGCGCAAGGTATATACAGCTCCCGATG 301  
|||  
Qy 81 AspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100  
|||  
Db 302 GATCAGTGAAGAGTACTGAGGGGACAATATCAAGCTATAGATGAGGCCATCAAGGAGAT 361  
|||  
Qy 101 Ala 101  
|||  
Db 362 GCG 364

RESULT 4  
D15350  
LOCUS  
DEFINITION RICC0499A Rice callus Oryza sativa (japonica cultivar-group) cDNA  
clone C0499, mRNA sequence.  
ACCESSION D15350 C98059  
VERSION D15350.2 GI:15072274  
KEYWORDS EST.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
1 (bases 1 to 530)  
Sasaki, T., Song, J., Koga-Ban, Y., Matsui, E., Fang, P., Higo, H.,  
Nagasaki, H., Hori, M., Miya, M. and Murayama-Kayano, E.  
Toward cataloging all rice genes: large-scale sequencing of

randomly chosen rice cDNAs from a callus cDNA library  
Plant J. 6 (4), 615-624 (1994)  
95078950  
MEDLINE  
PUBMED  
7987417  
COMMENT On May 17, 1993 this sequence version replaced gi:286543  
gi:3760805.  
Contact: Takuji Sasaki  
National Institute of Agrobiological Resources  
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
305-8602, Japan  
Tel: 81-298-38-7441  
Fax: 81-298-38-7468  
Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
PROJECT = RGP  
Seq primer: oligo(dT).  
Location/Qualifiers  
source  
1. .530  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="C0499"  
/clone\_lib="Rice callus"  
/note="Vector: pBluescript II SK+; Site 1: SalI; Site 2:  
NotI; cDNA prepared from rice callus mRNAs by using  
oligo(dT) as a primer and ligating to the SalI-NotI site  
of pBluescript II SK+ phagemid."

ORIGIN  
Alignment Scores:  
Pred. No.: 5.95e-47 Length: 530  
Score: 484.00 Matches: 96  
Percent Similarity: 96.00% Conservatives: 0  
Best Local Similarity: 96.00% Mismatches: 3  
Query Match: 91.32% Indels: 1  
DB: 7 Gaps: 0

US-10-629-953-4 (1-101) x D15350 (1-530)

Qy 1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGlyGluProAlaAlaLysArgArg 20  
|||  
Db 66 ATGNGCGGGAAGGGAACAGCGGTTCGGACGGCGCGCGCGCGCGCGCGCGCGCG 125  
|||  
Qy 20 gAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspIleValAlaGlnIle 40  
|||  
Db 126 TCCCGCGCGGACGACGCGCGCTCCGAGAGCGCGCGCGCGCGCGCGCGCGCGAT 185  
|||  
Qy 40 eSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValAlaAspIleArg 60  
|||  
Db 186 ATCGAAGAACAGGAGGTGGCGGTGGGACCTGGAAACGGCAAGGTGCTGTCGACATCG 245  
|||  
Qy 60 gGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMet 80  
|||  
Db 246 CGAGTCTACGAGAGGACGCGCAAGACCTCCCGCGCGCGCGCGCGCGCGCGCGAT 305  
|||  
Qy 80 tAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99  
|||  
Db 306 GGATCAGTGAAGATACCTGAGGGGACAATATCAAGCTATAGATGAGGCCATCAAGGAG 363  
|||

RESULT 5  
Call16904  
LOCUS  
DEFINITION SCAGLR1021A08.g LR1 Saccharum officinarum cDNA clone SCAGLR1021A08  
5', mRNA sequence.  
ACCESSION Call16904  
VERSION Call16904.1 GI:34970211  
KEYWORDS EST.  
SOURCE Saccharum officinarum  
ORGANISM Saccharum officinarum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum  
complex.



Dn 249 GATCAGTGGGAAGATACTACTGAGGACAAATATCAAGCCCATAGTAGGGCCATCAAGGAGAAC 308

RESULT 7  
CA230818 635 bp mRNA linear EST 25-SEP-2003  
LOCUS SCJFFL3C07F05.g Saccharum officinarum FL3 Saccharum officinarum  
DEFINITION cDNA Clone SCJFFL3C07F05 5', mRNA sequence.  
ACCSSION CA230818  
VERSION CA230818.1 GI:35294733  
KEYWORDS EST.  
SOURCE Saccharum officinarum  
ORGANISM Saccharum officinarum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum  
complex.

REFERENCE  
AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.  
TITLE The libraries that made SUCEST  
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
COMMENT Contact: Arruda P  
Centro de Biologia Molecular e Engenhariaia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089  
Email: parruda@unicamp.br  
Clone distribution: clone distribution information can be found  
through the Brazilian Clone Collection Center (BCCC) at  
<http://www.bccccenter.fcav.unesp.br>  
Plate: C07 row: F column: 05  
Seq primer: T7 Promoter Primer.  
Location/Qualifiers  
1 . 635  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCJFFL3C07F05"  
/lab\_host="DH108"  
/clone\_lib="Saccharum officinarum FL3"  
/note="Organ: Base of developing inflorescence (5cm-long); An  
vector: pSport1; Site 1: SalI; Site 2: NotI; An  
unidirectional cDNA library generated from [Base of  
developing inflorescence (5cm-long)]. cDNA was prepared  
from polyA+ mRNA using SuperScript Plasmid System Kit  
(Invitrogen). The double-strand cDNAs were fractionated  
in a sepharose CL-2B 40cm-columns and fragments sizing  
between 0.8 and 1.5 Kb were directionally cloned into the  
vector. Details of each source of RNA and library  
construction can be obtained at  
<http://sucest.lad.ic.unicamp.br/public>

ORIGIN

Alignment Scores:  
Pred. No.: 6,73e-40 Length: 635  
Score: 425.00 Matches: 79  
Percent Similarity: 89.00% Conservative: 10  
Best Local Similarity: 79.00% Mismatches: 11  
Query Match: 80.1% Indels: 0  
DB: 6 Gaps: 0

US-10-629-953-4 (1-101) x CA230818 (1-635)

Qy 1 MetTrpArglySGlyAenLyvArqPheGlyGlyVgluProAlaAlvArqArg 20  
Db 19 ATGTGGGGGAAGGAAAAGACGTGTTTCGGCGCGCGCGGCAGCGGCCGACAGCCGAC 78  
Qy 21 AlaAlaGlyAspGlyPyroSerGluSerAlaAspAspilleValValAlaGlnile 40  
Db 79 GCCCGCGGGATGACGGCCCCCTCCGAACCCCGCGAAGACCGTACGGTGCTAGCGGATA 138  
Qy 41 SerLyvAenArgvAlaAlaValArqThrTrpAenGlyLvsvValValAlaSpilleArg 60











```

ORGANISM  Hordeum vulgare subsp. vulgare
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
           Pooideae; Triticeae; Hordeum.
REFERENCE  1 (bases 1 to 579)
AUTHORS   Zhang,H., Weesche,W., Michalek,W., Stein,N. and Graner,A.
TITLE     EST sequencing and analysis in barley (2002)
JOURNAL   Unpublished (2002)
COMMENT   Molecular Markers Group, Department Genbank
           Institute of Plant Genetics and Crop Plant Research (IPK)
           Corrensstr. 3, 06466, Gatersleben, Germany
           Tel: 039482-5522
           Fax: 039482-5595
           Email: stein@ipk-gatersleben.de
           Insert Length: 579 Std Error: 0.00
           Plate: 12 row: 1 column: 13
           Seq primer: M13rev.
FEATURES  Location/Qualifiers
           source
             1..579
               /organism="Hordeum vulgare subsp. vulgare"
               /mol_type="mRNA"
               /cultivar="barke"
               /sub_species="vulgare"
               /db_xref="GABI:252770"
               /db_xref="taxon:112509"
               /clone="H112113"
               /tissue_type="female inflorescences"
               /dev_stage="female inflorescences (approx. 3 mm in size)"
               /lab_host="XLI10-Gold"
               /clone_lib="HI"
               /notes="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of
               cDNA); Site 2: XhoI (3'-end of cDNA); Due to a cloning
               artefact caused by the kit, in most cases the EcoRI site
               is NOT present, as well as the EcoRI adapter used for
               cloning. To excise the insert, restriction sites upstream
               EcoRI should be used (e.g. BamHI, Sali, PstI). NOTE: Also
               due to the cloning system used Blue/white selection for
               recombinants is not 100% reliable."
ORIGIN
Alignment Scores:
Pred. No.: 4,06e-38 Length: 579
Score: 409.50 Matches: 79
Percent Similarity: 88.00% Conservative: 9
Best Local Similarity: 79.00% Mismatches: 9
Query Match: 77.26% Indels: 3
DB: 5 Gaps: 1

US-10-629-953-4 (1-101) x BU998906 (1-579)
Qy 4 LysGlyValnLysArgPhe-----GlyGlyGlyGluProAlaAlaLysArgArg 20
Db 14 AAGGGGAACAAAGCGCTTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 73
Qy 21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspPileValAlaGlnIle 40
Db 74 GCGCGCGGGAAGGACGCGCCCTCCGAGGAACCGACGCGCATCGTCTCGCGCAGATA 133
Qy 41 SerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValAlaPileArg 60
Db 134 TCGAAGAACAAAGAGGGTGGCTGTGGAGGTGGTAATGCGAAGGTTCATGTCGATGCGC 193
Qy 61 GluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMet 80
Db 194 GAGTCTACGTGTCAGGAGCGGCAAGAGCTCCGACCCGCAAGAGGTATATCGCTCTCAATG 253
Qy 81 AspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100
Db 254 GATCAGTGGGAGATCTAGAGGACCAACATCAAGCCATAGACGAGCGCGTCAAGAGAAC 313

RESULT 15
CB860066/c

```

```

LOCUS     CB860066                               618 bp    mRNA    linear    EST 22-APR-2003
DEFINITION Hi12113w HI Hordeum vulgare subsp. vulgare cDNA clone Hi12113
3-PRIME, mRNA sequence.
ACCESSION CB860066
VERSION   CB860066.1 GI:30054625
KEYWORDS  EST.
SOURCE    Hordeum vulgare subsp. vulgare
ORGANISM  Hordeum vulgare subsp. vulgare
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
           Pooideae; Triticeae; Hordeum.
REFERENCE  1 (bases 1 to 618)
AUTHORS   Zhang,H., Weesche,W., Michalek,W., Stein,N. and Graner,A.
TITLE     EST sequencing and analysis in barley (2002)
JOURNAL   Unpublished (2002)
COMMENT   Contact: Stein Nils
           Molecular Markers Group, Department Genbank
           Institute of Plant Genetics and Crop Plant Research (IPK)
           Corrensstr. 3, 06466, Gatersleben, Germany
           Tel: 039482-5522
           Fax: 039482-5595
           Email: stein@ipk-gatersleben.de
           Insert Length: 618 Std Error: 0.00
           Plate: 12 row: 1 column: 13
           Seq primer: T7.
FEATURES  Location/Qualifiers
           source
             1..618
               /organism="Hordeum vulgare subsp. vulgare"
               /mol_type="mRNA"
               /cultivar="barke"
               /sub_species="vulgare"
               /db_xref="GABI:557649"
               /db_xref="taxon:112509"
               /clone="H112113"
               /tissue_type="female inflorescences"
               /dev_stage="female inflorescences (approx. 3 mm in size)"
               /lab_host="XLI10-Gold"
               /clone_lib="HI"
               /notes="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of
               cDNA); Site 2: XhoI (3'-end of cDNA); Due to a cloning
               artefact caused by the kit, in most cases the EcoRI site
               is NOT present, as well as the EcoRI adapter used for
               cloning. To excise the insert, restriction sites upstream
               EcoRI should be used (e.g. BamHI, Sali, PstI). NOTE: Also
               due to the cloning system used Blue/white selection for
               recombinants is not 100% reliable."
ORIGIN
Alignment Scores:
Pred. No.: 4,38e-38 Length: 618
Score: 409.50 Matches: 79
Percent Similarity: 88.00% Conservative: 9
Best Local Similarity: 79.00% Mismatches: 9
Query Match: 77.26% Indels: 3
DB: 6 Gaps: 1

US-10-629-953-4 (1-101) x CB860066 (1-618)
Qy 4 LysGlyValnLysArgPhe-----GlyGlyGlyGluProAlaAlaLysArgArg 20
Db 607 AAGGGGAACAAAGCGCTTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCAG 548
Qy 21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspPileValAlaGlnIle 40
Db 547 GCGCGCGGGAAGGACGCGCCCTCCGAGGAACCGACGCGCATCGTCTCGCGCAGATA 488
Qy 41 SerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValAlaPileArg 60
Db 487 TCGAAGAACAAAGAGGGTGGCTGTGGAGGTGGTAATGCGAAGGTTCATGTCGATGCGC 428
Qy 61 GluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMet 80
Db 427 GAGTCTACGTGTCAGGAGCGGCAAGAGCTCCGACCCGCAAGAGGTATATCGCTCTCAATG 368

```

Qy 81 AspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100  
Db 367 GATCAGTGGGAGACTGTAGGGACACATCAAGCCATAGACGAGGCCGTCNAGGAGAAC 308

Search completed: January 12, 2005, 03:27:53  
Job time : 2553 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: January 12, 2005, 01:19:56 ; Search time 2548 Seconds  
(without alignments)  
1444.432 Million cell updates/sec

Title: US-10-629-953-4

Perfect score: 530  
Sequence: 1 MRRKGNRRFGGGGRRPAARR.....QWTLIRDNIAKIDAEIKENA 101

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool.h/US10629953/runat.10012005.172536.9425/app.query.fasta.1.263  
-DB=EST -OPMT=fasted -SUPPLX=p2n.ref -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bites -START=1 -END=1 -MATRIX=blomsum62 -TRANS=unman40.cdi -LIST=45  
-DOCLIGN=200 -THR\_SCORE=Pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTWTPCTO -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10629953 @cgn 1.1 3437 @runat.10012005.172536.9425 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: gb\_eest1:  
2: gb\_eest2:  
3: gb\_hrc:  
4: gb\_eest3:  
5: gb\_eest4:  
6: gb\_eest5:  
7: gb\_eest6:  
8: gb\_ges1:  
9: gb\_ges2:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	530	100.0	697	6	CB684473 OSJNEFI4A
2	530	100.0	784	6	CB684474 OSJNEFI4A
3	517	97.5	693	1	AU082888 AU082888
4	484	91.3	530	7	D15350 R1C0499A.R
5	425	80.2	523	6	CA116904 SCAGRI102
6	425	80.2	631	6	CA151445 SCJFRZ201
7	425	80.2	635	6	CA230818 SCJFLT300
8	425	80.2	637	6	CA259070 SCBORT101
9	425	80.2	649	6	CA189554 SCCCR100

10	416	78.5	440	7	CF760736	CF760736 DSAF1.65
11	415.5	78.4	564	6	CD863080	CD863080 A201.105K
12	415	78.3	855	6	CA129813	CA129813 SCUTR201
13	412.5	77.8	532	5	BQ838824	BQ838824 WMS3593_B
14	409.5	77.3	579	5	BQ938906	BQ938906 H112133
15	409.5	77.3	618	6	CB860066	CB860066 H112133w
16	409.5	77.3	618	6	CB860077	CB860077 H112J02w
17	409.5	77.3	619	5	BQ938918	BQ938918 H112J02r
18	409.5	77.3	1072	2	CK212042	CK212042 FCAS02390
19	400.5	75.6	539	1	AU082741	AU082741 AU092741
20	398	75.1	568	6	CB250145	CB250145 3529_1_18
21	398	75.1	596	6	CD568912	CD568912 3529_1_19
22	396	74.7	470	6	CD913769	CD913769 G550.119C
23	395	74.5	603	1	AI691445	AI691445 606018A12
24	392	74.0	606	6	CA189705	CA189705 SCCCR100
25	391	73.8	403	2	BE424113	BE424113 WHE0076.C
26	391	73.8	618	1	AI734589	AI734589 606033E08
27	390	73.6	566	2	BE186543	BE186543 946008A01
28	389.5	73.5	460	6	CA710341	CA710341 WDK2C.PK0
29	381	71.9	623	7	CF635823	CF635823 zmrw00.0
30	375	70.8	516	6	CF019517	CF019517 OBN1d08_x
31	373.5	70.5	586	6	CD568913	CD568913 3529_1_99
32	363	68.5	588	7	CF625939	CF625939 zmrw805_0
33	363	68.5	588	7	CF627159	CF627159 zmrw805_0
34	363	68.5	652	3	AY112536	AY112536 Zea mayb
35	362.5	68.4	443	6	C28075	C28075 C28075 Rice
36	347.5	65.6	396	6	C27852	C27852 C27852 Rice
37	338.5	63.9	555	4	BI795760	BI795760 H030F04.E
38	329	62.1	288	1	AU068341	AU068341 AU068341
39	329	62.1	542	5	BQ498860	BQ498860 946170D05
40	299	56.4	485	1	AI943769	AI943769 618045D08
41	272.5	51.4	586	6	CB079577	CB079577 hp73903.b
42	272	51.3	551	6	CB922664	CB922664 VVD08E01
43	272	51.3	570	5	BQ796180	BQ796180 EST 5118
44	272	51.3	601	6	CB339739	CB339739 CA33E1031
45	272	51.3	620	6	CB921948	CB921948 VVD077E10

## ALIGNMENTS

RESULT 1  
LOCUS CB684473 697 bp mRNA linear EST 09-APR-2003  
DEFINITION OSJNEFI4A12.f OSJNEF Oryza sativa (japonica cultivar-group)  
clone OSJNEFI4A12 5', mRNA sequence.

ACCESSION CB684473  
VERSION CB684473.1 GI:29688198  
KEYWORDS EST.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
1 (bases 1 to 697)  
Jantanasuriyarat.C., Lu.G., Gowda.M., Hatfield.J., Zhou.B., Mazur.E.,  
Kudrna.D., Dean.R., Soderlund.C., Wang.R. and Wang.G.  
Large-scale identification of ESTs involved in the interaction  
between rice and Magnaporthe grisea  
Unpublished (2003)  
JOURNAL  
COMMENT  
Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
85721-0088, USA  
Tel: 520 626 3967  
Fax: 520 621 9288  
Email: http://genome.arizona.edu  
PCR Primers  
FORWARD: gta aac cga cga cca gtc  
BACKWARD: gga aac agc tat gac cat g  
Plate: 14 row: A column: 12  
Seq primer: gta aac cga cga cca gtc.  
Location/Qualifiers

## FEATURES

**THIS PAGE BLANK (uspto)**

**THIS PAGE BLANK (uspto)**

## COMMENT

Contact: Takuji Sasaki  
National Institute of Agrobiological Resources  
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
305-8602, Japan  
Tel: 81-298-38-7441  
Fax: 81-298-38-7468  
Email: tsasaki@abrr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
PROJECT = "RGP"

## FEATURES

## source

1. 693

Location/Qualifiers

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="C53889"

/note="Vector: pBluescript II SK+ Site 1: SalI; Site 2:

NotI; cDNA prepared from rice callus mRNA by using

oligo(dT) as a primer and ligating to the SalI-NotI site

of pBluescript II SK+ phagemid."

## ORIGIN

## Alignment Scores:

Pred. No.: 1,056-50 Length: 693  
Score: 517.00 Matches: 100  
Percent Similarity: 99.01% Conservative: 0  
Best Local Similarity: 99.01% Mismatches: 1  
Query Match: 97.55% Indels: 0  
DB: 1 Gaps: 0

US-10-629-953-4 (1-101) x AU082888 (1-693)

## ORIGIN

## Alignment Scores:

Pred. No.: 5,956-47 Length: 530  
Score: 484.00 Matches: 96  
Percent Similarity: 96.00% Conservative: 0  
Best Local Similarity: 96.00% Mismatches: 3  
Query Match: 91.32% Indels: 1  
DB: 2 Gaps: 0

US-10-629-953-4 (1-101) x D15350 (1-530)

JOURNAL  
MEDLINE  
PUBMED  
COMMENT  
On May 17, 1993 this sequence version replaced gi:286543  
gi:3760805.

Contact: Takuji Sasaki  
National Institute of Agrobiological Resources  
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
305-8602, Japan  
Tel: 81-298-38-7441  
Fax: 81-298-38-7468  
Email: tsasaki@abrr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
PROJECT = "RGP"

Seq primer: oligo(dT).  
Location/Qualifiers

## source

1. 530

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="C0499"

/note="Vector: pBluescript II SK+ Site 1: SalI; Site 2:

NotI; cDNA prepared from rice callus mRNA by using

oligo(dT) as a primer and ligating to the SalI-NotI site

of pBluescript II SK+ phagemid."

Randomly chosen rice cDNAs from a callus cDNA library

Plant J. 6 (4), 615-624 (1994)

95078950

7987417

95078950

95078950

On May 17, 1993 this sequence version replaced gi:286543  
gi:3760805.

Contact: Takuji Sasaki  
National Institute of Agrobiological Resources  
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
305-8602, Japan  
Tel: 81-298-38-7441  
Fax: 81-298-38-7468  
Email: tsasaki@abrr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
PROJECT = "RGP"

Seq primer: oligo(dT).  
Location/Qualifiers

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="C0499"

/note="Vector: pBluescript II SK+ Site 1: SalI; Site 2:

NotI; cDNA prepared from rice callus mRNA by using

oligo(dT) as a primer and ligating to the SalI-NotI site

of pBluescript II SK+ phagemid."

## COMMENT

Contact: Takuji Sasaki  
National Institute of Agrobiological Resources  
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
305-8602, Japan  
Tel: 81-298-38-7441  
Fax: 81-298-38-7468  
Email: tsasaki@abrr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
PROJECT = "RGP"

## FEATURES

## source

1. 693

Location/Qualifiers

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="C53889"

/note="Vector: pBluescript II SK+ Site 1: SalI; Site 2:

NotI; cDNA prepared from rice callus mRNA by using

oligo(dT) as a primer and ligating to the SalI-NotI site

of pBluescript II SK+ phagemid."

## ORIGIN

## Alignment Scores:

Pred. No.: 1,056-50 Length: 693  
Score: 517.00 Matches: 100  
Percent Similarity: 99.01% Conservative: 0  
Best Local Similarity: 99.01% Mismatches: 1  
Query Match: 97.55% Indels: 0  
DB: 1 Gaps: 0

US-10-629-953-4 (1-101) x AU082888 (1-693)

## ORIGIN

## Alignment Scores:

Pred. No.: 5,956-47 Length: 530  
Score: 484.00 Matches: 96  
Percent Similarity: 96.00% Conservative: 0  
Best Local Similarity: 96.00% Mismatches: 3  
Query Match: 91.32% Indels: 1  
DB: 2 Gaps: 0

US-10-629-953-4 (1-101) x D15350 (1-530)

JOURNAL  
MEDLINE  
PUBMED  
COMMENT  
On May 17, 1993 this sequence version replaced gi:286543  
gi:3760805.

Contact: Takuji Sasaki  
National Institute of Agrobiological Resources  
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
305-8602, Japan  
Tel: 81-298-38-7441  
Fax: 81-298-38-7468  
Email: tsasaki@abrr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
PROJECT = "RGP"

Seq primer: oligo(dT).  
Location/Qualifiers

## source

1. 530

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="C0499"

/note="Vector: pBluescript II SK+ Site 1: SalI; Site 2:

NotI; cDNA prepared from rice callus mRNA by using

oligo(dT) as a primer and ligating to the SalI-NotI site

of pBluescript II SK+ phagemid."

## COMMENT

Contact: Takuji Sasaki  
National Institute of Agrobiological Resources  
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
305-8602, Japan  
Tel: 81-298-38-7441  
Fax: 81-298-38-7468  
Email: tsasaki@abrr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
PROJECT = "RGP"

## FEATURES

## source

1. 693

Location/Qualifiers

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="C53889"

/note="Vector: pBluescript II SK+ Site 1: SalI; Site 2:

NotI; cDNA prepared from rice callus mRNA by using

oligo(dT) as a primer and ligating to the SalI-NotI site

of pBluescript II SK+ phagemid."

## ORIGIN

## Alignment Scores:

Pred. No.: 1,056-50 Length: 693  
Score: 517.00 Matches: 100  
Percent Similarity: 99.01% Conservative: 0  
Best Local Similarity: 99.01% Mismatches: 1  
Query Match: 97.55% Indels: 0  
DB: 1 Gaps: 0

US-10-629-953-4 (1-101) x AU082888 (1-693)

## ORIGIN

## Alignment Scores:

Pred. No.: 5,956-47 Length: 530  
Score: 484.00 Matches: 96  
Percent Similarity: 96.00% Conservative: 0  
Best Local Similarity: 96.00% Mismatches: 3  
Query Match: 91.32% Indels: 1  
DB: 2 Gaps: 0

US-10-629-953-4 (1-101) x D15350 (1-530)

JOURNAL  
MEDLINE  
PUBMED  
COMMENT  
On May 17, 1993 this sequence version replaced gi:286543  
gi:3760805.

Contact: Takuji Sasaki  
National Institute of Agrobiological Resources  
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
305-8602, Japan  
Tel: 81-298-38-7441  
Fax: 81-298-38-7468  
Email: tsasaki@abrr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
PROJECT = "RGP"

Seq primer: oligo(dT).  
Location/Qualifiers

## source

1. 530

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="C0499"

/note="Vector: pBluescript II SK+ Site 1: SalI; Site 2:

NotI; cDNA prepared from rice callus mRNA by using

oligo(dT) as a primer and ligating to the SalI-NotI site

of pBluescript II SK+ phagemid."

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**